University of Alberta

Exposure Misclassification in Reproductive Epidemiology Studies Investigating Disinfection By-products in Drinking Water

by

Karina A.M. Bodo

©

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment
of the requirements for the degree of Doctor of Philosophy

in

Medical Sciences–Public Health Sciences

Department of Public Health Sciences

Edmonton, Alberta
Spring, 2002

National Library of Canada
Acquisitions and Bibliographic Services
395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale du Canada
Acquisitions et services bibliographiques
395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-68548-9

Canada

University of Alberta

Library Release Form

Name of Author: Karina A.M. Bodo

Title of Thesis: Exposure Misclassification in Reproductive Epidemiology Studies Investigating Disinfection By-products in Drinking Water

Degree: Doctor of Philosophy

Year this Degree Granted: 2002

Permission is hereby granted to the University of Alberta Library to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only.

The author reserves all other publication and other rights in association with the copyright in the thesis, and except as herein before provided, neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatever without the author's prior written permission.

Karina Bodo

10960-75th Ave.
Edmonton, Alberta, Canada T6G 0G9

Date submitted: April 16, 2002

University of Alberta

Faculty of Graduate Studies and Research

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled **Exposure Misclassification in Reproductive Epidemiology Studies Investigating Disinfection By-products in Drinking Water** submitted by Karina Aranyka Michelle Bodo in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Medical Sciences–Public Health Sciences.

Dr. Steve E. Hrudey

Dr. Kenneth L. Froese

Dr. A. Senthilselvan

Dr. Chris X. Le

B. Kratochvil
Dr. Byron Kratochvil

Dr. Tong Yu

L D Saunders
For Dr. Mark Nieuwenhuijsen

Date approved: 2 April 2002

"It is the nature of life that we lose face before we find wisdom, fall to our knees before we look up to the heavens, and face our darkness before we see the light. Each of us wanders through the wilderness of experience to gather worldly wisdom. Everything we encounter serves in its own way. We succeed by failing, learn by our mistakes, and rise to great heights by a winding staircase."

-Dan Millman

To Simon

Abstract

The challenges of developing accurate exposure assessment for epidemiology studies investigating a possible causal association between exposure to disinfection by-products (DBP) in treated drinking water and adverse reproductive and developmental effects are discussed. Exposure assessment in these studies can and should be improved. But, how good does exposure assessment have to be for these epidemiology studies? Various aspects of this question are investigated in this dissertation. A framework was developed to group past and future studies according to their exposure assessment. It was observed that few epidemiology studies to date employed the higher levels of exposure assessment. Additionally, exposure assessment in these epidemiology studies has not improved over time. Future studies should strive to measure specific agents and obtain personal data. The potential for exposure misclassification from a variety of factors was investigated and the effect of that exposure misclassification on the odds ratio (OR) of a hypothetical epidemiology study was quantified. Ultimately, the degree to which exposure assessment must be taken depends on several key factors, including the DBP species measured, the type of disinfection, the frequency of sampling or monitoring, and the type of data treatment. It was found that under specific circumstances, principally found in the chloraminated system investigated here, exposure assessment of the type used in some epidemiology studies is not predicted to result in exposure misclassification serious enough to have a substantial effect on the resulting OR, particularly at low predicted "true" ORs. However, the observations as a whole confirm that the current situation of measuring unknown causal agents at sampling points that are spatially and temporally removed from the individual study subjects is unsatisfactory. As expected, the

observations suggest that the measurement of known causal agents close to the point of exposure and during critical exposure time periods is preferable, particularly at higher expected true ORs. The potential for exposure misclassification and OR attenuation introduced by categorizing continuous DBP concentration data were investigated. The categorization of inadequate exposure data can not replace adequate and representative continuous exposure data. Future studies should carefully weight the effects of using categorical rather than continuous data.

Acknowledgements

I would like to thank my supervisors, Dr. Steve Hrudey and Dr. Kenneth Froese for their financial and intellectual support. In addition, I would like to thank the members of my supervisory committee, in particular Dr. A. (Sentil) Senthilselvan for his encouragement and collaboration.

I would also like to thank the water utilities in Edmonton, Calgary, and Winnipeg for generously letting me use their monitoring data.

In addition, I would like to express my gratitude to those who helped along the journey:

In the lab at the University of Alberta: Linda Kimpe, Jeff Rose, Mike Ongley, Barbara Ells, Shirley Samuel

At EPCOR Water Services: Simon Thomas, Les Gammie, Lloyd Penner

At Calgary Water Works and Calgary sampling volunteers: Trevor Satchwill, Darcy Petkau, Terry Hallet, John Jagorinec, Leanne Satchwill, Hugh Costello, Alan Beairsto, Heinz and Elfriede Wenger, Irene Wenger, Chris Ouellette

Edmonton sampling volunteers: Kim Arndt, Michael Ongley, Cindy Jardine, Errol Billing, Simon Thomas, Irene Wenger, Chris Ouellette

Outside the University of Alberta: Dr. Barry Thomas, Dr. Will King

I would like to thank my family and the many friends who have been my cheering squad and who have picked up the pieces when needed. You were always there to wipe away the tears and to celebrate the victories. I feel incredibly fortunate that there are too many of you to list by name. However, I would particularly like to thank my parents for their unwavering love and support and for always believing in me. Lastly and most importantly: Thank you, Simon. My love always.

Table of Contents

Page

Chapter 1: Introduction 1

Criteria of causality 2

Types of epidemiology studies 4

Factors affecting exposure assessment 6

DBP formation 7

Toxicology 12

Exposure pathways 17

Biomarkers 26

References 30

Chapter 2: Exposure Assessment Framework 39

Personal information vs. Aggregated population information 39

Specific vs. non-specific chemicals 40

Direct measurement vs. indirect measurement 40

Summary of Epidemiology Studies 56

Discussion 75

Conclusions 87

References 89

Chapter 3: Misclassification 93

Introduction 93

What is misclassification? 93

Epidemiological measures of strength of association 95

Indices of accuracy of measurement 95

Effect of misclassification on odds ratios 98

Development of Test Model 120

Applying the methods for predicting the effect of exposure misclassification on the strength of association of epidemiology studies 121

Page

Categorical treatment of data 122
Continuous treatment of data 124
References 126

Chapter 4: Data Sources and Data Collection 128
Monitoring data 128
Description of water utilities 129
    City A 130
    City B 130
    City C 131
Experimental Data 131
    Daily sampling for DBPs in drinking water 131
    Summary statistics for experimental and monitoring data 134
    K_{ow} experiments 142
References 144

Chapter 5: Potential for Exposure Misclassification and Effect on OR of Variations in DBP Concentrations between the Water Treatment Plant and Distribution System 145
Introduction 145
Analysis of continuous exposure data 147
Analysis of categorical exposure data 160
Summary and Conclusions 167
References 169

Page

Chapter 6: Potential for Exposure Misclassification and Subsequent OR Attenuation from Measuring TCM or TTHM rather than More Toxicologically Plausible Causal Agents 170
Introduction 170
Analysis of continuous exposure data 171
Application of method to literature values 179
Analysis of categorical exposure data 180
Summary and conclusions 186
References 186

Chapter 7: Potential for Exposure Misclassification and Subsequent OR Attenuation as a Result of Monthly Convenience Sampling 187
Introduction 187
Analysis of continuous exposure data 188
Analysis of categorical exposure data 193
Conclusions 195

Chapter 8: Effects of categorical vs. continuous treatment of exposure data 197

Chapter 9: Summary and Conclusions 211
Bibliography 221
Appendices
Appendix for Chapter 3: Logistic regression primer and Tables 3-6 and 3-10 234
Appendix for Chapter 4: Raw data, MDLs, and DCAN correction factor 275
Appendix for Chapters 5, 6, and 7: Tables of correlations, sensitivites, specificities, and Calculated ORs from Chapters 5, 6, and 7 333

List of Tables

Table 1-1. Selected chlorine and chloramine disinfection by-products

Table 1-2. Doses from DBP toxicology studies and corresponding adult human doses

Table 2-1. Summary of epidemiology studies investigating a link between drinking water and various adverse reproductive effects, sorted by exposure category

Table 2-2. Summary of Agents of Exposure

Table 2-3. Summary of OR medians and ranges

Table 3-1. Example two-by-two table for the calculation of relative risk and odds ratio

Table 3-2. Example of a table to determine the sensitivity and specificity of a study

Table 3-3. Effects of non-differential exposure misclassification on dichotomous exposure variables (Rothman and Greenland, 1998)

Table 3-4. Calculated "true" ORs for the Klotz and Pyrch (1999) study at different levels of sensitivity and specificity[a]

Table 3-5. Calculated "true" ORs for the Bove et al. (1992) study at different levels of sensitivity and specificity

Table 3-6. Effect of non-differential misclassification of categorical data

Table 3-7. Selected data on effects of misclassification on categorical data, $P_N = 0.2$

Table 3-8. Effects of non-differential exposure misclassification on polychotomous exposure variables (Rothman and Greenland, 1998)

Table 3-9. Another example of the possible effects of non-differential misclassification of polychotomous exposure variables (Wacholder et al., 1991)

Table 3-10. Effects of non-differential misclassification on continuous data

Table 3-11. Selected data on effects of misclassification on continuous data

Table 3-12. Sensitivity and specificity table of urine excretion rate vs. ingestion dose

Table 3-13. Sensitivity and specificity table of urine excretion rate vs. water concentration

Table 4-1. City A monitoring data

Table 4-2. City B monitoring data

Table 4-3. City C monitoring data

Table 4-4. City A daily summary statistics

Table 4-5. City B daily summary statistics

Table 4-6. City A monthly summary statistics

Table 4-7. City B monthly summary statistics

Table 4-8. City C monthly summary statistics

Table 4-9. Experimental and Calculated K_{ow} Values for Selected Chlorinated DBPs

Table 5-1. Correlation between daily DBP concentration data at water treatment plant #1 and distribution system locations in City A, and its effects on the odds ratio

Table 5-2. Summary of low and high correlations for water treatment plant data compared to home monitoring and distribution system data

Table 5-3. Number of data sets with r-values resulting in measured ORs (OR_X) within 10% of the true ORs (OR_T)

Table 5-4. Sensitivity and specificity between daily DBP concentration data at water treatment plant #1 and distribution system locations in City A, and their effects on the odds ratio

Table 5-5. Number of measured ORs within 10% of the true OR

Table 6-1. Correlation between TCM and other DBP species at water treatment plants 1 and 2 and distribution system locations in City A, and its effects on the odds ratio

Table 6-2. Summary of high and low r-values for investigating measured causal agent

Table 6-3. Number of r-values within the required range

Table 6-4. OR attenuation method applied to correlations between DBP species from the literature

Table 6-5. Sensitivity and specificity between TCM and other DBP concentrations at water treatment plants 1 and 2 and distribution system locations in City A, and their effects on the odds ratio

Table 6-6. Number of measured ORs within 10% of the true OR

Table 7-1. Water treatment plant #1 TCM data in City A; correlations between the "measured" convenience sampling data and the "true" data with the resulting OR attenuation due to exposure misclassification

Table 7-2. Number of data set pairs with correlation coefficients within the required range for the OR_X to be within 10% of the OR_T

Table 7-3. Categorical treatment of City A data comparing the "measured" convenience sampling data with the "true" values and the resulting effect on the OR

Table 7-4. Number of data set pairs with measured ORs within 10% of the true OR

Table 8-1. Correlations between daily DBP concentration data at water treatment plant #1 and home sampling locations in City A, and their effects on the odds ratio: Continuous data

Table 8-2. Sensitivity and specificity between daily DBP concentration data at Water treatment plant #1 and distribution system locations in City A, and their effects on the odds ratio: Halves

Table 8-3. Sensitivity and specificity between daily DBP concentration data at Water treatment plant #1 and distribution system locations in City A, and their effects on the odds ratio: Tertiles

Table 8-4. Sensitivity and specificity between daily DBP concentration data at Water treatment plant #1 and distribution system locations in City A, and their effects on the odds ratio: Quartiles

Table 8-5. Sensitivity and specificity between daily DBP concentration data at Water treatment plant #1 and distribution system locations in City A, and their effects on the odds ratio: Sextiles

Table 8-6. Example of a 2x2 table to determine the sensitivity and specificity of a study

Table 8-7. Three examples of the influence of specificity and sensitivity on OR attenuation

Table 8-8. Summary of inconsistency between categorical and continuous treatment of the data

In the Appendices:

Table 3-6. Effect of non-differential misclassification of categorical data

Table 3-10. Effects of non-differential misclassification on continuous data

Raw data

Method detection limits (MDLs)

Chapter 5 continuous data tables

Chapter 5 categorical data tables

Chapter 6 continuous data tables

Chapter 6 categorical data tables

Chapter 7 data tables

List of Figures

Figure 2-1. Framework for Categories of Exposure Assessment Used in Reproductive Epidemiological Studies

Figure 2-2. Results of epidemiology studies investigating stillbirths

Figure 2-3. Results of epidemiology studies investigating low birth weights

Figure 2-4. Results of epidemiology studies investigating neural tube defects

Figure 2-5. Results of epidemiology studies investigating spontaneous abortion

Figure 3-1. Effect of measurement error on the distribution of a normally distributed exposure, after Armstrong et al., 1992.

Figure 3-2. Effect of sensitivity (specificity = 0.8, $P_N = 0.2$)

Figure 3-3. Effect of sensitivity (specificity = 0.99, $P_N = 0.2$)

Figure 3-4. Effect of specificity (sensitivity = 0.8, $P_N = 0.2$)

Figure 3-5. Effect of specificity (sensitivity = 0.99, $P_N = 0.2$)

Figure 3-6. Effect of non-differential measurement error on the exposure distributions of the non-diseased and diseased groups, after Armstrong et al., 1992

Figure 3-7. Effect of non-differential measurement error on the odds ratio curve, after Armstrong et al., 1992

Figure 3-8: Continuous summary: ρ vs. OR

Figure 3-9: Continuous summary: ρ^2 vs. OR

Figure 5-1. OR attenuation as a results of exposure misclassification:Effect of r

Figure 5-2. % change in OR as a result of exposure misclassification

Figure 5-3. Percent of data set pairs with measured ORs within 10% of the true OR (OR_T) for daily sampling, continuous treatment

Figure 5-4. Percent of data set pairs with measured ORs within 10% of the true OR (OR_T) for monthly sampling, continuous treatment

Figure 5-5. Percent of data set pairs with measured ORs within 10% of the true OR (OR_T) for daily sampling, categorical treatment

Figure 5-6. Percent of data set pairs with measured ORs within 10% of the true OR (OR_T) for monthly sampling, categorical treatment

Figure 6-1. Percent of data set pairs with measured ORs within 10% of the true OR (OR_T) for daily sampling, continuous treatment

Figure 6-2. Percent of data set pairs with measured ORs within 10% of the true OR (OR_T) for monthly sampling, continuous treatment

Figure 6-3. Percent of data set pairs with measured ORs within 10% of the true OR (OR_T) for monthly sampling, continuous treatment

Figure 6-4. Percent of data set pairs with measured ORs within 10% of the true OR (OR_T) for daily sampling, categorical treatment

Figure 6-5. Percent of data set pairs with measured ORs within 10% of the true OR (OR_T) for monthly sampling, categorical treatment

Figure 6-6. Percent of data set pairs with measured ORs within 10% of the true OR (OR_T) for monthly sampling, categorical treatment

Figure 7-1. Percent of data set pairs with measured ORs within 10% of the true OR (OR_T), continuous treatment

Figure 7-2. Percent of data set pairs with measured ORs within 10% of the true OR (OR_T), categorical treatment

Figure 8-1. Anomalous example showing the effect of sensitivity and specificity on OR_X

Figure 8-2. Example of the effect of sensitivity and specificity on OR_X

Figure 8-3. Example of the effect of sensitivity and specificity on OR_X

In Appendix for Chapter 4

Sample chromatogram: 10 μg/L

List of Abbreviations and Glossary

BCAA: bromochloroacetic acid

BCAN: bromochloroacetonitrile

BDCAA: bromodichloroacetic acid

BDCM: bromodichloromethane

CDBM: chlorodibromomethane

CH: chloral hydrate

CI: confidence interval

DBAA: dibromoacetic acid

DBAN: dibromoacetonitrile

DBCM: dibromochloromethane

DBP: disinfection by-product

DCA: dichloroacetaldehyde

DCAA: dichloroacetic acid

DCAN: dichloroacetonitrile

DOC: dissolved organic carbon

HAA: haloacetic acid

HAA$_5$: The sum of five of the haloacetic acids (chloroacetic acid, dichloroacetic acid, trichloroacetic acid, bromoacetic acid, dibromoacetic acid)

HAN: haloacetonitrile

MBAA: monobromoacetic acid

MCAA: monochloroacetic acid

NOM: natural organic matter

NRC: National Research Council

SUVA: specific UV absorbance; (UV absorbance at 254 nm x 100) / (DOC concentration)

Tald: total aldehydes

TBAA: trobromoacetic acid

TBM: tribromomethane; bromoform

TCAA: trichloroacetic acid

TCAN: trichloroacetonitrile

TCM: trichloromethane; chloroform

TCNM: trichloronitromethane

THAA: total haloacetic acids

THAN: total haloacetonitriles

THM: trihalomethane

TOC: total organic carbon

TTHM: total trihalomethanes

US EPA: United States Environmental Protection Agency

VOC: volatile organic compounds

WHO: World Health Organization

Accuracy: the correctness of an experimental result, subject to both random and systematic errors[1].

Bias:

- *In epidemiology:* when the mean value of an imperfect measurement differs systematically from the true mean value, the measurement technique is said to be biased[2]; types of bias include, but are not restricted to, measurement error, selection bias, observation bias, interviewer bias, confounding, and recall bias.
- *In chemistry:* similar to the epidemiology definition, bias occurs when the mean value of the imperfect measurement differs from the true mean value as a result of systematic errors. Types of systematic errors include instrument errors, operator errors, and method errors[1].

Categorical data: a means for classifying data whereby the observed value is replaced by a category representing a range of values into which the observed value falls, used as a method of viewing exposure or outcome data in epidemiology; data are divided into categories such as exposed/not exposed (dichotomous), or 0-10 µg/L, 11-20 µg/L, 21-30 µg/L, 31-40 µg/L, 41-50 µg/L (polychotomous).

Coefficient of determination (r^2): The extent to which the variability in one measure may be accounted for (or predicted) through knowledge of the value of the other measure[3].

Combined chlorine residual: the chloramine concentration in the water after the reaction with organics during the disinfection step in the water treatment plant[4].

Continuous data: retention of the original observed value of the data, used as a method of viewing exposure or outcome data in epidemiology; the measurement of data is taken on a continuous scale.

Dichotomous data: see *Categorical data*

Differential exposure misclassification: errors occur differently in the classification of subjects as exposed or not exposed depending on whether the subjects have the outcome or not.

Fetotoxic: having a harmful or poisonous effect on a fetus as the result of exposure to a toxin.

Free chlorine residual: the sum of the chlorine, hypochlorous acid, and hypochlorite ion concentrations in the water after the reaction with organics during the disinfection step in the water treatment plant.

Gavage: forced feeding by a stomach tube[5].

Misclassification of exposure (exposure misclassification): the wrong measure of exposure is assigned to epidemiology study subjects e.g. study subjects are classified as exposed when they are in actuality not exposed.

Nondifferential exposure misclassification: errors occur in the classification of subjects as exposed or not exposed with the same frequency regardless of whether the subjects have the outcome or not.

Odds ratio (OR): The ratio of the odds of exposure among the cases to the odds of exposure among the controls.

Pearson product moment correlation coefficient (r): a measure of the strength of the linear relationship between two variables[3].

Polychotomous data: see *Categorical data*

Precision:

- *In epidemiology:* the lack of random error; precision can be increased by 1) increasing the study size and 2) modifying the study design to improve the efficiency of information collection from the study subjects[6].
- *In chemistry:* the reproducibility of results; the degree of mutual agreement among data obtained the same way[1].

Precursors: the organic compounds in untreated water that contribute to the formation of chlorinated disinfection by-products upon reaction with chlorine-based disinfectants[4].

Relative risk (RR): The ratio of the incidence of disease in a group of exposed subjects to the incidence of disease in a group of nonexposed subjects.

Sensitivity:

- *In epidemiology:* the proportion of those who truly are exposed who are classified as being exposed.
- *In chemistry:* the ability of an instrument or a method to discriminate between small differences in analyte concentration[1].

Specificity: In epidemiology, the proportion of those who are truly not exposed who are classified as being not exposed.

Teratogenesis: the origin or mode of production of a malformed fetus; the disturbed growth processes involved in the production of a malformed neonate[5].

References

1. Skoog, D.A. and J.J. Leary, 1992. *Principles of Instrumental Analysis*. Montreal: Saunders College Publishing.

2. Kelsey, J.L., W.D. Thompson, and A.S. Evans, 1986. *Methods in Observational Epidemiology*. Monographs in Epidemiology and Biostatistics. A.M. Lilienfeld, ed. Vol. 10. New York: Oxford University Press. 366.

3. Dawson-Saunders, B. and R.G. Trapp, 1994. *Basic and Clinical Biostatistics*. 2nd ed. Norwalk, Connecticut: Appleton and Lange.

4. Sawyer, C.N., P.L. McCarty, and G.F. Parkin, 1994. *Chemistry for Environmental Engineering*. 4th ed. New York: McGraw-Hill, Inc. 658.

5. Spraycar, M., ed. 1995. *Stedman's Medical Dictionary*. 26th ed. Williams and Wilkins: Baltimore.

6. Rothman, K.J. and S. Greenland, 1998. *Modern Epidemiology*. 2nd ed. Philadelphia: Lippincott-Raven Publishers.

Chapter 1
Introduction

In the early 1970's, it was discovered that the combination of chlorine and natural organic matter in the water produced halogenated organic compounds. In 1974, chloroform was identified as one of these compounds (Rook, 1974). A subsequent survey conducted in the United States determined that chlorinated by-products were a major contaminant of chlorinated surface water (Morris et al., 1992). This finding instigated a flurry of epidemiology studies aimed at determining the possible health effects of chronic, widespread exposure to these chemicals in drinking water.

A variety of adverse health effects have been investigated by these epidemiologic studies. Initial studies concentrated on cancer outcomes, with bladder cancer providing the most consistent association (Mills et al., 1998; Cantor et al., 1987; Freedman et al., 1997; King and Marrett, 1996; McGeehin et al., 1993; Zierler et al., 1988; Gottleib et al., 1982; Young et al., 1981; Wilkins and Comstock, 1981; Brenniman et al., 1980; Alvanja et al., 1978).

A possible link between chlorinated drinking water and adverse reproductive effects was first identified in a study investigating a suspected cluster of adverse pregnancy outcomes in relation to water contamination in California (Deane et al., 1989). This study was inconclusive with respect to a relationship between the suspected contaminant (trichloroethane) and birth outcomes; however, a higher than average rate of adverse pregnancy outcomes in women consuming large quantities of water was observed. This observation led to the investigation of other substances in chlorinated drinking water as a possible cause. Since then, many studies have investigated a possible link between treated drinking water or specific chemicals in treated drinking water and adverse reproductive effects, with inconsistent results. A recent study (Graves et al., 2001) investigated the weight of evidence of all toxicology and epidemiology studies that have looked at a possible association between DBPs and adverse reproductive or developmental effects. This study looked at many outcomes with respect to exposure to DBPs and concluded that there is either no evidence of association[1], the results are mixed, weak or inconsistent[2], or the results are suggestive of positive outcomes[3]. A major shortcoming of these studies has been exposure assessment (Reif et al., 1996; Nieuwenhuijsen et al., 2000, Graves et al., 2001), particularly with regards to individual exposure assessment.

---

[1] Low birth weight, very low birth weight, preterm delivery, caesarean delivery, congenital anomalies, spina bifida, cleft lip and palate, cardiac anomalies, gastrointestinal anomalies, genital anomalies, integument anomalies, musculoskeletal anomalies, chromosomal abnormalities, and neonatal death

[2] Neonatal jaundice, all congenital anomalies/birth defects, all central nervous system anomalies, neural tube defects, respiratory anomalies, spontaneous abortion/miscarriage, stillbirth/fetal death

[3] Growth retardation including term low birth weight, intrauterine growth retardation, or small for gestational age, small body length, small cranial circumference, urinary tract defects.

This section will discuss the difficulties in assessing individual exposure to chlorinated DBPs in drinking water, as well as the importance of exposure assessment in evaluating a causal relationship between an effect and an outcome. Reproductive epidemiology studies completed to date will be described in detail with respect to their exposure assessment for possible causal agents in chlorinated drinking water. A framework will be introduced with which these epidemiology studies can be compared on the basis of their exposure assessment. The framework will be employed to answer several questions:

-How well have reproductive epidemiology studies to date characterized individual exposure to possible causal agents in treated drinking water and have there been useful improvements in exposure assessment over time?

-Can improved exposure assessment clarify a possible causal association between exposure to agents in treated drinking water and adverse reproductive effects?

-How can future epidemiology studies improve exposure assessment in the absence of biomarkers of exposure?

Criteria of causality

Criteria of causality provide a valuable reference in elucidating the importance of accurate individual exposure assessment in epidemiologic studies (Beaglehole et al., 1993). The guidelines for judging causation presented here are derived from a set of concepts set out by Sir Austin Bradford Hill and by the U.S. Surgeon General (Hill, 1965; U.S. Public Health Service, 1964). All of the proposed causal criteria will be outlined here with emphasis on those criteria for which exposure assessment is particularly important. These concepts and the subsequent guidelines provide an approach by which a potential relationship between an exposure and an effect may be judged to be causal or not. These guidelines are discussed here in the order in which they should be considered when examining a potential cause-effect relationship.

The first criterion that must be considered when examining an exposure-outcome relationship for causality is the *temporal relationship*. Simply stated, the cause must precede the effect. If the cause does not precede the effect, then there are no grounds upon which to base causality. This criterion demands an accurate reckoning of the time relationship between the proposed exposure and the resulting outcome and as such requires a certain level of accuracy in the determination of both the exposure and the outcome.

The second criterion to be considered is *plausibility*: is it biologically plausible that the exposure will cause the expected outcome? This question is best answered by toxicology studies. An exposure-outcome relationship with a biologically plausible mechanism provides a strong argument in favour of causality.[4]

---

[4] If a biologically plausible mechanism is not obvious, the argument for causality is not negated. Biological plausibility is often dependent on the state of the science at the time of investigation. If

The third criterion is *consistency*: if several different studies with a variety of designs, carried out in different locations and under different conditions, consistently report the same result then the argument for causality is strengthened. However, a lack of consistency does not necessarily preclude a causal association. The differing study designs and circumstances (such as exposure levels) could reduce the impact of the causal agent in some of the studies. Therefore the studies with the best designs must be given the most weight when evaluating this criterion.

The fourth criterion is the *dose-response relationship*. If the frequency or severity of the outcome increases with increasing frequency or magnitude of exposure to the potential causal agent, then a dose-response relationship is seen. In an unbiased study, a clear dose-response relationship can be a good indication of a causal relationship. For those exposures and effects for which a dose-response relationship is valid, defining that dose-response relationship depends on defining both the dose and the response accurately. It follows that to have confidence in a dose-response relationship, there must be accuracy in the identification and quantitation of the exposure, as well as in the determination of the outcome.

The fifth criterion is the *strength of the association*, measured by the risk ratio of the study. A large risk ratio argues more strongly for causality than does a small one.[5] In this fourth criterion, the necessity of accurate individual exposure assessment is emphasized. It is widely accepted that if an association is causal, weak exposure assessment resulting in non-differential misclassification of exposure will bias the resulting risk ratio towards the null value, suggesting a weaker association than is actually true. The logic follows, therefore, that if an association is causal, more accurate exposure assessment will increase the apparent strength of the association. There is always the possibility that a weak association really is an indication of a non-causal association. In this case all efforts must be made to develop a strong study design and accurate exposure assessment to ensure an unbiased study so that the conclusion of a non-causal association can be made with confidence.

The sixth criterion is *reversibility*. Reversibility states that if a potential causal agent is removed, the likelihood of the outcome occurring is decreased. If the removal of a potential causal agent results in a decrease in the occurrence of the outcome, then the argument for a causal relationship is strengthened. It is necessary in evaluating this criterion that an accurate assessment of exposure or non-exposure to the potential causal agent be determined. It is important to note that this criterion is limited to causal processes that involve reversible

---

biological plausibility is not evident at the time of the investigation of the causal relationship, it may become apparent in the future.

[5] However, a small risk ratio does not preclude a causal association since the size of the risk ratio can depend on the prevalence of other possible causes in relation to the agent of interest.

mechanisms, which is not the case with adverse reproductive and developmental outcomes.

The seventh criterion is *study design*. There are many different epidemiologic study designs, each with a different level of ability to test for causation. The best study design for testing causation is the randomized controlled trial or clinical trial[6]. Cohort studies and case-control studies are the next-best study designs and are commonly used in environmental epidemiology. Cross-sectional studies are the least able to test causation because they usually do not give evidence of a temporal relationship (the cause must precede the effect). The various study designs are outlined in a subsequent section of this thesis.

A caveat to the previous seven criteria is necessary. There is really no hard and fast rule for judging a potential causal relationship. All available evidence must be taken into account when determining whether an exposure really is the cause of an outcome. In the case of conflicting evidence, the types of evidence must be weighted with respect to their relative positions within the criteria for causality. Most importantly, a temporal relationship between the cause and the effect must be established. If the effect precedes the alleged cause, then there is no argument for a causal relationship. Plausibility, consistency, and dose-response are the next most important criteria. The likelihood of a causal association is increased if evidence representing several of the criteria all points to the same conclusion.

Types of Epidemiology Studies

As mentioned in the seventh of the criteria for causality, there are several types of epidemiology study designs with varying abilities to test causality. A description of the different study designs is presented here to emphasize the importance of the study design to the ability of an epidemiologic study to test for causality, with respect to the criteria for causality. In particular, it should be noted that the study designs designated as better at testing causality in these causality criteria all use exposure and outcome data for individuals rather than for populations. The utility of a study in testing a hypothesis of causality depends in part on whether individual exposure can be linked to individual outcome. It follows, then, that the more accurately one can characterize the exposure and the outcome in each individual, the more useful a study will be in testing a hypothesis of causality.

There are two basic categories under which epidemiology study designs fall: *experimental*, and *observational or non-experimental* (Beaglehole et al., 1993; Rothman and Greenland, 1998)[7]. In experimental studies the investigator assigns the exposure levels and follows subjects for subsequent changes in health status[8]. Experimental study designs will not be explored in depth here because the reproductive epidemiology studies looking at exposure to DBPs are all observational studies.

---

[6] This type of study is rarely done in environmental epidemiology because of the ethical limits on experimentation with humans.

[7] It should be noted here that the classification of study designs as "experimental", "non-experimental", "observational", "analytic", and "descriptive" involves semantics. These categories and the words used to describe them are not consistent throughout the field of epidemiology and different texts will group studies differently (for an example of this, compare Beaglehole et al., 1993 and Rothman and Greenland, 1998 described here with Hennekens and Buring, 1987).

Observational or non-experimental studies fall under two categories of study design: *analytical studies* that investigate a relationship between health status and other variables, and *descriptive studies* that simply describe the health status of a community based on information already available[9]. Descriptive studies do not compare health status in relation to other factors. The reproductive epidemiology studies investigated here all fall under the category of analytical studies. There are several types of analytical study designs including cohort studies, case-control studies, cross-sectional studies, and aggregated studies[10].

Aggregated studies use data from whole populations to compare disease patterns between different groups within a population during the same period of time or to compare disease patterns among the same group over several time periods. The units of observation are populations rather than individuals. Aggregated studies tend to be relatively quick and inexpensive to conduct as the information required is often already available from public records. Aggregated studies are often a first step in investigating a possible relationship between an exposure and a disease. However, there is a major disadvantage in aggregated studies that limits their usefulness. Because aggregated studies use data for the whole population rather than for individuals, exposure cannot be linked to disease in individuals. This can lead to a phenomenon called the "ecological fallacy"[11] when inappropriate conclusions are made regarding relationships between exposures and outcomes based on aggregated data from populations rather than individuals. This arises because population rate data do not allow any determination of whether individuals who experience the outcome were also exposed. Aggregated studies are generally used to propose epidemiologic hypotheses, not to test them.

Cross sectional studies, also called prevalence studies, measure disease state and exposure in individuals in a population at the same point in time. This type of study provides a "snapshot" of the state of a population with respect to specific exposures and diseases at any one particular time. The major limiting factor of

---

However, the description of individual study designs (cohort, case-control, cross-sectional, and ecological) and their weaknesses and strengths are, in general, now consistent throughout the field.

[8] Types of experimental studies include randomized controlled trials (also called clinical trials), field trials, and community intervention and cluster randomized trials. Randomized control trials are rated the most useful study design for testing causation. However, this type of study is rarely used in environmental epidemiology because of the ethical problems associated with experimentally exposing subjects to potentially harmful agents.

[9] Usually from public data bases

[10] Aggregated studies have been referred to in the past as ecological studies; however, this terminology is changing to reflect more accurately the description of these studies as aggregating information over populations rather individuals. "Ecological studies" is a misleading term since these studies have nothing to do with "ecology".

[11] From "ecological studies". See footnote 7.

cross-sectional studies is that it is usually unclear whether exposure preceded or followed the health outcome. This does not make cross-sectional studies generally useful for testing epidemiologic hypotheses.[12]

In *case control studies*, subjects are selected based on whether they do (cases) or do not (controls) have the health outcome in question. The groups are then compared with respect to the proportion of each group with the exposure or characteristic of interest. Case control studies are relatively inexpensive and take less time to complete relative to cohort studies. They offer a solution to the difficulties of studying health outcomes with long latency periods, and they allow an investigation of many etiologic exposures or characteristics for a specific health outcome. One disadvantage results from the fact that both the exposure and the disease must have occurred at the start of the study. This fact subjects case control studies to possible selection bias of cases *vs.* controls based on exposure status, or differential reporting of exposure data based on disease status.

In *cohort studies*, study groups are designated according to exposure status (exposed or not exposed). At the time of the exposure, all subjects must be free of the disease. Subjects are then compared based on the proportions of exposed and non-exposed individuals who develop the outcome of interest subsequent to the exposure.[13] The exposed and non-exposed subjects are then followed for occurrence of the health outcomes of interest. Because study subjects are free from the disease at the time of initiation of the study, the temporal sequence between the exposure and the outcome can be established. In addition, because the study groups are selected based on exposure status, cohort studies are ideal for studying rare exposures or for studying multiple outcomes from the same exposure. However, cohort studies are very time-consuming and expensive. As subjects must be followed for many years after exposure, there is the potential for bias due to losses of subjects to follow-up, particularly in prospective cohort studies.

Factors Affecting Exposure Assessment

As mentioned previously, measurement of individual exposure has been a limitation of many of the epidemiology studies investigating a relationship between agents in chlorinated drinking water and adverse reproductive effects. There are several factors that contribute to the determination of individual exposure to agents in chlorinated drinking water. One factor is the formation of DBPs during chlorination and in the time between chlorination and exposure of the individual, which affects the concentrations of DBPs at the point of exposure.

---

[12] For the special circumstance when the values of the exposure variables do not change over time, meaning that the current exposure values are the same as the exposure values at the initiation of the health outcome, cross sectional studies may be used to test epidemiological hypotheses. However, this does not entirely resolve the issue of temporality of exposure and outcome.

[13] Cohort studies can be retrospective or prospective. In retrospective studies all exposures and outcomes have occurred at the initiation of the study. Exposure status is determined from a time before the outcomes occurred. In prospective studies, the exposure may or may not have occurred at the initiation of the study, but the outcomes have certainly not occurred.

Second, there is the issue of which agent is the causal agent[14]. Third is individual exposure, which varies depending on the pathways of exposure and the route of uptake into the body. These three main factors are described in detail in the following sections.

DBP formation

The concentration of DBPs is a significant factor in the determination of human exposure to DBPs in drinking water. For example, ingestion dose can be calculated as the volume of drinking water ingested multiplied by the DBP concentration in that water. However, there are many factors contributing to the variability of DBP concentrations in treated drinking water. These factors are discussed below.

To date there have been dozens of halogenated and non-halogenated disinfection by-products identified in chlorinated drinking water (Krasner, 1999b; Richardson, 1998). The most abundant group of DBPs by weight are the trihalomethanes (THMs), followed by the haloacetic acids (HAAs)[15]. A list of several known chlorination and chloramination by-products is found in Table 1-1.

---

[14] The majority of epidemiology studies that have investigated a particular agent have concentrated their efforts on the trihalomethane (THM) group of compounds. However, toxicology studies have investigated other groups of disinfection by-products (DBPs) also found in chlorinated drinking water.

[15] There are several alternative disinfectants in use today, such as chlorine dioxide, chloramine, and ozone, and these disinfectants produce their own suites of disinfection by-products. Only DBPs created by the reaction of organic matter with chlorine or chloramine will be discussed here. Disinfection by-products produced as a result of chloramination are basically the same as those produced by chlorination, with the exception that cyanogen chloride is preferentially found in chloraminated water (Krasner et al., 1989).

Table 1-1. Selected chlorine and chloramine disinfection by-products

DBP Group	Individual DBP
Trihalomethanes (THMs)	chloroform (trichloromethane, TCM), bromodichloromethane (BDCM), chlorodibromomethane (CDBM), bromoform (tribromomethane, TBM)
Haloacetic Acids (HAAs)	monochloroacetic acid (MCAA), dichloroacetic acid (DCAA), trichloroacetic acid (TCAA), monobromoacetic acid (MBAA), dibromoacetic acid (DBAA), tribromoacetic acid (TBAA), bromochloroacetic acid (BCAA), bromodichloroacetic acid (BDCAA), chlorodibromoacetic acid (CDBAA)
Haloacetonitriles (HANs)	dichloroacetonitrile (DCAN), trichloroacetonitrile (TCAN), dibromoacetonitrile (DBAN), bromochloroacetonitrile (BCAN)
Haloketones (HKs)	1,1-dichloro-2-propanone, 1,1,1-trichloro-2-propanone
Aldehydes	formaldehyde, acetaldehyde
Others	chloral hydrate (CH, also trichloroacetaldehyde), chloropicrin (trichloronitromethane), cyanogen chloride

To understand the formation mechanics of chlorinated DBPs, it is important to understand something of the chemistry of chlorine and chloramine. Gaseous chlorine ($Cl_{2(g)}$) quickly dissolves when added to water and establishes a series of equilibrium reactions with hypochlorous acid (HOCl) and hypochlorite ion (OCl^-) (Snoeyink and Jenkins, 1980).[16]

$$Cl_{2(g)} \leftrightarrow Cl_{2(aq)}; \qquad K_{sp} = 6.2 \times 10^{-2}$$
$$Cl_2 + H_2O \leftrightarrow HOCl + H^+ + Cl^-; \qquad K_{eq} = 4 \times 10^{-4} \text{ at } 25\ ^{\circ}C$$
$$HOCl \leftrightarrow H^+ + OCl^-; \quad pK_a = 7.5 \text{ at } 25\ ^{\circ}C$$

The effectiveness of a disinfectant depends on the nature of the disinfectant, the concentration of the disinfectant, and the contact time. A strong disinfectant may take less time or lower concentrations to achieve the same disinfecting action as higher concentrations of a weaker disinfectant used over a longer time period. Chlorine is a strong oxidizing agent and therefore a powerful disinfectant. Because of this property, chlorine tends to react very quickly, thereby requiring higher concentrations of chlorine to be added to water to ensure a chlorine residual over long periods of time or distances in a distribution system. To stabilize the chlorine oxidation reaction and decrease the amount of chlorine required to maintain a residual, ammonia is sometimes added to chlorinated water to form chloramines.[17] The reactions of chlorine and ammonia to form chloramines are as follows:

---

[16] In addition, hypochlorous acid can react with bromide ions in the water to form hypobromous acid (HOBr) following the reaction: $Br^- + HOCl \rightarrow HOBr + Cl^-$. Note that this reaction is not reversible and that HOBr is more reactive than HOCl, which has particular implications for DBP formation. Bromide ion is naturally found in some raw waters. Bromine (Br_2) is sometimes found as a contaminant in the Cl_2 used in disinfection.

$$NH_3 + HOCl \rightarrow NH_2Cl + H_2O \text{ (monochloramine)}$$
$$NH_2Cl + HOCl \rightarrow NHCl_2 + H_2O \text{ (dichloramine)}$$
$$NHCl_2 + HOCl \rightarrow NCl_3 + H_2O \text{ (trichloramine)}$$

Monochloramine and dichloramine are better disinfectants than trichloroamines and are the desired products of the ammonia-chlorine reaction[18].
The general equation for the formation of DBPs during chlorination is:

$$HOCl + NOM \rightarrow \text{halogenated DBPs}$$

Where:
NOM is the natural organic matter in the raw water[19].

Other factors affecting the formation of DBPs at the water treatment plant in addition to the amount and nature of *NOM* are *pH*, *temperature*, *contact time*, and *chlorine dose*. One study suggested that pH has a very large effect on DBP formation (Stevens et al., 1989). In pilot plant experiments, the pH of the disinfection reaction and contact time were varied, and the temperature, chlorine dose, and organic content of the water were kept constant. This study found that the formation of THMs increased with increasing pH over time. DCAA and TCAA represented the HAA group. TCAA showed decreased formation over time as pH increased whereas DCAA seemed unaffected by changes in pH.[20] In general, the concentrations of DBPs increase with increasing contact time; however this increase is pH dependent to some extent.

---

[17] Chloramines are themselves strong disinfectants; however, they are not as strong disinfectants as chlorine and therefore tend to be more stable over time. As a result, some water treatment plants utilize chlorine as a primary disinfectant to provide the initial disinfecting action and add ammonia to produce chloramines, which provide a more stable chlorine residual in the distribution system. Chlorine residuals from chlorine, hypochlorous acid, and hypochlorite ions are referred to as free chlorine residuals. Chlorine residuals from chloramines are referred to as combined chlorine residuals.

[18] The extent of halogenation of the ammonia can be controlled with attention to the pH of the reaction and the mole ratio of chlorine to ammonia (Sawyer et al., 1994).

[19] The nature of NOM is one of the factors affecting the type and amount of DBPs formed during chlorination. NOM is a mixture of humic and non-humic substances. The humic substances are a mixture of humic and fulvic acids and are complex organic compounds with unknown structures that leach from decaying vegetation. NOM is considered to be a precursor of DBP formation. Humic substances generally have a higher DBP formation potential than the non-humic fraction. Surrogate measures of the humic content of the water and of the DBP precursors present in the water are total organic carbon (TOC) and specific UV absorbance (SUVA). SUVA is calculated as the UV absorbance at 254 nm multiplied by 100 and divided by the concentration of the dissolved fraction of organic carbon in the water (DOC) (Krasner, 1999a; Krasner, 1999b).

[20] The formation of chloral hydrate increased over time at pH 5 and 7, but decreased over time at pH 9.4. The formation of DCAN, the representative of the HANs group, increased over time at pH 5, decreased at pH 7 and occurred at low levels at pH 9.4. Looking at the effect of contact time on the formation of brominated vs. chlorinated species of THMs, the study found that TCM formation increased over time; however, TBM increased rapidly and then remained constant. The

The temperature of the raw water and the chlorine dose are important factors in the rate of DBP formation. These factors, as well as the nature and amount of NOM, vary depending on the season. The temperature of the raw water tends to be higher in the summer, which speeds up the rate of formation of DBPs. This in turn mandates higher chlorine doses and residuals. High chlorine doses and residuals favour the formation of HAAs over THMs, and of trihalogenated HAAs over mono- or dihalogenated HAAs (Krasner, 1999b). A factor contributing to variations in DBP speciation is the seasonal variation in bromide ion in the raw water. The amount of bromide ion in the water is influenced by rainfall and runoff patterns, salt water intrusion into source water during low flow and high demand conditions, inflow from agricultural drainage, and the blending of raw waters of different composition to make up for a potential shortfall during high demand periods. These factors are seasonally dependent (Krasner, 1999b).

Similar factors[21] govern the formation of DBPs in chloraminated water. In general, it is presumed that once ammonia is added to chlorinated water all DBP formation ceases (Singer, 1999); however, this is not always the case[22].

The DBP concentration of the water leaving the water treatment plant depends on the characteristics of the raw water[23], the treatment processes and chemical doses employed in the water treatment plant[24], and the contact time[25]. However, the formation of DBPs does not cease when the treated water leaves the water treatment plant. Indeed, to provide some protection and to act as a real-time indicator of contamination occurring in the distribution system, a chlorine residual is maintained in the distribution system. This residual is available not only for disinfection, but also for reaction with organic matter in the distribution system[26] and resultant DBP formation.

---

mixed brominated/chlorinated species (BDCM and CDBM) showed similar behaviour to TBM. The authors of the study postulated that the difference between TBM and TCM may be a result of the differing DBP formation kinetics of HOBr and HOCl, with HOBr reactions occurring more quickly than HOCl reactions.

[21] pH, temperature, amount and nature of NOM, bromide ion concentration and contact time.

[22] A few studies were undertaken to investigate DBP formation upon disinfection with chloramine only, with no prior treatment with chlorine. It was found that more DBPs are produced at low pH and high chlorine-to-nitrogen ratios during chloramination. In addition, bromide ions in the source water caused a decrease in the production of dichloramine and an increase in brominated DBP species (Diehl et al., 1995). A subsequent study found that higher pHs and low chlorine-to-nitrogen ratios resulted in a decrease in DBP formation. It was also found that dihalogenated acid species were preferentially formed during chloramination as compared to the preferential formation of trihalogenated acid species during chlorination (Diehl et al., 2000).

[23] Amount and type of NOM, bromide ion concentration, and temperature

[24] pH, disinfectant type and dose, residual type and concentrations

[25] The flow rates and the residence time in the water treatment plant

[26] The organic matter is mostly provided from NOM that has precipitated on the pipe walls forming a biofilm sludge which can exert a significant chlorine demand (Rossman et al., 1999; Valentine et al., 1999).

The formation of DBPs in the distribution system is dependent on many of the same factors as DBP formation in the water treatment plant. Higher temperatures in the distribution system will increase the rate of formation of DBPs, temperature being largely dependent on season. High pH favours THM formation, while low pH favours HAA formation, with the dihalogenated acids insensitive to pH (Singer, 1999). The contact time to the point of use (usually, the resident's home) also has a large effect on the concentrations of DBPs at the point of use, since DBP formation continues with contact time[27]. The type and concentration of residual affects the formation of disinfection byproducts in the distribution system. Free chlorine residual will continue to form THMs and HAAs; however, DBP formation is expected to virtually cease upon the addition of ammonia to create combined chlorine residuals (Singer, 1999). In addition, the higher the free chlorine concentration in the treated water, the faster the DBP formation. The formation of different DBPs species in the distribution system can vary depending on reaction kinetics (Singer, 1999)[28].

The formation of DBPs in the distribution system is dependent on the chemical characteristics of the water entering the system[29], the nature of the distribution system itself[30], and the individual DBP species in question. As a result of these factors, there can be significant temporal and spatial variability in DBP species and concentrations at various points in the distribution system. These variations can occur on a seasonal, monthly, daily, or even hourly basis (Singer, 1999). A study of water consumption and water use activities of pregnant women and their partners found that while the majority of tap water (67%) was consumed at home, the balance (33%) was consumed in other places[31] (Shimokura et al., 1998). Another study found that cold tap water consumed at home accounted for 39 ± 17% of total liquid consumption, while cold tap water at work accounted for 7 ± 9% (Froese et al., 2002, In Press). Given the variations in DBP concentrations from location to location in the distribution system, this information on water consumption patterns has significant implications for assessing exposure to DBPs at the individual level.

Changes in the concentrations of disinfection by-products do not cease upon arrival of the water at a consumer's residence. There are many water use activities in a home that can affect the concentrations of DBPs in the water. Water is heated in water heaters and for cooking or beverage making. It has been found that the formation reaction of chloroform will go to completion after storage for some time in a hot water tank (Benoit, 1999). Several studies have found that heating and boiling water for the purposes of making food or beverages affects the DBP concentrations in the water (Lahl et al., 1982[32]; Benoit et al., 2000[33]). Heated and cold water are both used to prepare beverages and food. Bearing in mind the goal of water utilities to maintain a chlorine residual to the consumer's tap, it is likely that the addition of this chlorinated water to organic matter in food and beverages could result in the formation of additional DBPs (Balko et al., 2001[34]). Other studies have shown that the method of storage of water (in an open or closed container, in the fridge, on the counter), water temperature, and the use of filters affect the concentrations of DBPs in water in the consumer's home[35].

---

[27] Contact time in the distribution system is the time it takes for the water to travel from the water treatment plant to the consumer's place of residence. The contact time will depend on the location of the consumer with respect to the water treatment plant, the water demand patterns in the distribution system, the design of the distribution network, the storage volume in the system, and water production at the water treatment plant (Singer, 1999).

[28] The kinetics of formation of HAAs are faster than those of THMs. Therefore, more HAA formation will occur in the water treatment plant than in the distribution system, whereas much THM formation will occur in the distribution system. In addition, some species, particularly the HANs and the HKs, will decay in the distribution system due to continuing reaction with chlorine.

[29] pH, residual type and concentration, temperature

[30] Contact time and amount of organic matter available for reaction

[31] For example, at work, at a restaurant, at a friend's home, at a relative's home

Toxicology

As with the epidemiology studies, toxicology studies initially concentrated on the potential carcinogenicity of DBP compounds. More recently, the reproductive toxicity of DBPs has been investigated. While reproductive epidemiology studies have been limited to an investigation of THMs for the most part, reproductive toxicology studies have investigated several groups of DBP compounds, including THMs, HAAs, and HANs. These studies were done on a variety of animals, including rats, mice, rabbits, and dogs. For all these toxicology studies, the doses administered to the animals correspond to extremely high doses in humans. A summary and comparison of the doses can be found in Table 1-2.

THMs

Toxicology studies have been carried out to investigate the reproductive effects of THM via both ingestion and inhalation. Ingestion studies found decreased maternal weight gain (Thompson et al., 1974[36]; Ruddick et al., 1983[37]; Narotsky et al., 1997[38]), reduced birth weight (Thompson et al., 1974; Ruddick et al., 1983), reduced fetal weight (Ruddick et al., 1983), interparietal effects (Ruddick et al., 1983), and full litter reabsorption (Narotsky et al., 1997). Studies on inhaled chloroform in air found increased litter reabsorption (Schwetz et al., 1974[39]; Murray et al., 1979[40]), retarded fetal development (Schwetz et al., 1974), fetal deformities (Murray et al., 1979), and abnormal spermatozoa (Land et al., 1981[41]).

---

[32] It was found that the chloroform levels decreased by 50% after heating for 1 minute at 80 °C, by 70% immediately after the water began to boil, by 80% after boiling for 3 minutes, and by 90% after boiling for 5 minutes.

[33] In this study by Health Canada it was found that boiling water in a kettle or pot for 2 or 5 minutes reduced HAA concentrations by 4% whereas concentrations of the remaining 21 DBPs in the water, including THMs, HANs, haloketones (HKs), chloral hydrate, chloropicrin, and cyanogen chloride, were reduced by greater than 96%.

[34] A study was carried out at the University of Alberta investigating TCAA and DCAA concentrations in coffee made with cold water containing residual chlorine. TCAA and DCAA concentrations were elevated in the coffee compared to the initial cold water. In addition, as chlorine residual in the cold water increased in successive trials, the TCAA and DCAA concentrations in the coffee increased proportionally.

[35] It has been observed that approximately 20% of THMs will volatilize from the water during the time it takes to draw the water from a tap into a glass and drink it (Wallace, 1997). A Health Canada study found that while virtually all THMs and HANs are removed from water stored at room temperature in an open container for 8 days, only 21% of HAAs are removed under the same conditions (Benoit et al., 2000). The same Health Canada study also found that both pressure filters and drip filters decreased the concentrations of all the chlorinated DBPs in the water.

[36] Thompson and colleagues investigated the teratological effect of chloroform on rats and rabbits (Thompson et al., 1974). Rats were dosed with 0, 20, 50, or 126 mg/kg/day on days 6-15 of

HAAs

HAA toxicology studies involved ingestion exposure exclusively. Adverse reproductive outcomes included decreased maternal weight gain (Smith et al., 1989a[42]; Smith et al., 1992[43];), increased litter reabsorption (Smith et al., 1989a),

---

gestation. Rabbits were dosed with 0, 20, 35, or 50 mg/kg/day on gestational days 6-18. Both species showed decreased maternal weight gain. Both species also exhibited reduced birth weight at the highest doses, but no teratological effects were observed.

[37] Ruddick and colleagues administered each of the four THMs separately by gavage on gestational days 6-15 to groups of pregnant Sprague-Dawley rats (Ruddick et al., 1983). The doses administered were 100, 200, and 400 mg/kg/day for TCM and 50, 100, and 200 mg/kg/day for TBM, BDCM, and CDBM. They found maternal effects such as decreased maternal weight gain in the groups administered TCM and the highest doses of BDCM and CDBM, and liver enlargement in the groups administered TCM. In addition, a decreased fetal body weight was observed in the groups administered TCM. Other fetotoxic effects such as interparietal anomalies were observed in the groups administered TBM and BDCM. A dose-response relationship was not observed during the treatment and no teratogenic effects were observed.

[38] In investigating the effects of the gavage vehicle (water vs. corn oil) on the developmental toxicity of BDCM on pregnant Fischer-344 rats, Narotsky and colleagues (Narotsky et al., 1997) observed decreases in maternal weight gain and increased full-litter reabsorption upon exposure to BDCM via either vehicle. Rats were dosed between gestational days 6-15 with 0, 25, 50, or 75 mg/kg/day. Effects on maternal weight gain were more pronounced in the aqueous vehicles at lower doses. Full-litter reabsorption was observed in both vehicles at the two highest doses.

[39] Schwetz and colleagues conducted a study on the effects of inhaled chloroform in air on fetal rats (Schwetz et al., 1974). Pregnant Sprague-Dawley rats were exposed to 30, 100, or 300 ppm$_v$ of chloroform in air for 7 hr/day over gestational days 6 through 15. A high incidence of fetal reabsorption was seen at the highest dose. Fetal development was retarded at all doses and decreased fetal body measurements were seen at 30 and 300 ppm$_v$. The study authors concluded that inhaled chloroform is embryotoxic, but not teratogenic to rats.

[40] Murray and colleagues investigated the effects of inhaled chloroform in air on mice by dosing pregnant CF-1 mice with 0 or 100 ppm$_v$ of chloroform in air for 7 hr/day on gestational days 1-7, 6-15, or 8-15 (Murray et al., 1979). The ability to maintain pregnancy was significantly impaired in mice exposed on days 1-7 or 6-15, but no teratogenic effects were observed. The fetuses of mice exposed on days 8-15 showed an increased incidence of cleft palate. All groups showed decreased ossification of bones. Decreased fetal body measurements were seen in the groups dosed on days 1-7 and 8-15.

[41] In order to investigate the effect of inhaled chloroform in air on spermatozoa in mice, Land and colleagues dosed the animals at 0.08% (by volume) and 0.04% (by volume) chloroform in air for 4 hr/day over 5 days (Land et al., 1981). Twenty-eight days after exposure the epididymal spermatozoa were examined. Abnormal spermatozoa were observed in mice exposed to both doses of chloroform compared to controls. The authors suggest that inhalation exposure to chloroform is genetically toxic.

[42] In their study investigating the teratogenic effects of TCAA, Smith and colleagues administered TCAA to pregnant Long-Evans rats (Smith et al., 1989a). TCAA was administered by gavage at different concentrations (1, 330, 800, 1200 and 1800 mg/kg/day) in water during gestational days

fetal developmental defects or malformations (Smith et al., 1989a; Smith et al., 1992; Epstein et al., 1992[44]), and decreased fetal growth (Smith et al., 1989a). Male reproductive effects were also studied and some adverse effects were found (Cicmanec et al., 1991[45]; Toth et al, 1992[46]; Linder et al., 1994[47]).

HANs

Toxicology studies on ingestion of HANs also found reproductive and developmental effects. Some of the effects found included full litter reabsorption

---

6-15. Maternal effects observed included decreased weight gain at the higher doses and increased spleen and kidney weights proportional to dose. At the higher doses, there was a dose-response relationship in terms the percent of reabsorbed implants in each litter. Other dose-dependent effects observed were decreased weight and length in live fetuses, and cardiovascular system soft tissue defects in fetuses. This study concluded that TCAA is teratogenic in rats at doses 330 mg/kg/day and greater.

[43] Smith and colleagues carried out two studies on pregnant Long-Evans rats investigating the developmental toxicity of DCAA (Smith et al., 1992). Rats were orally dosed with doses of 0, 900, 1400, 1900, or 2400 mg/kg/day in water in one study and 0, 14, 140, or 400 mg/kg/day in water in the other study during days 6-15 of gestation. Adverse maternal effects observed included deaths at the three highest doses (1400, 1900 and 2400 mg/kg/day), a reduction in maternal weight gain at all except the lowest doses, and a dose-related increase in the weights of liver, spleen and kidneys. Fetal adverse effects included soft-tissue effects, particularly in the cardiovascular system, that were more frequent at higher doses.

[44] Epstein and colleagues investigated the cardiopathic effects of DCAA on fetal Long-Evans rats (Epstein et al., 1992). In a series of tests, pregnant rats were dosed with 1) 1900 mg/kg/day on gestational days 6-8, 9-11 or 12-15; 2) 2400 mg/kg/day on gestational days 10, 11, 12, or 13; 3) 3500 mg/kg/day on gestational days 9, 10, 11, 12, or 13; 4) 1900 mg/kg/day on gestational days 6-15. For the first group dose with 1900 mg/kg/day, heart malformations were seen in the fetuses of those rats dosed on gestational days 9-11 and 12-15. Doses of 2400 mg/kg/day resulted in heart malformations when dosed on days 10 and 12, whereas doses of 3500 mg/kg/day produced heart malformations when dosed on gestational days 9, 10, and 12.

[45] Cicmanec and colleagues studied the chronic toxicity of DCAA over a 90 day period on male and female dogs (Cicmanec et al., 1991). Doses were administered orally via gelatin capsules at 0, 12.5, 39.5, and 72 mg/kg/day. Pregnant dogs were not used in this study so teratogenic effects were not observed. However, many other adverse effects were observed involving the brain, testicles, lungs, pancreas, and liver in most of the high-dose dogs and some of the mid- and low-dose dogs. The observed effect to the testicles of the male dogs has implications for the reproductive toxicity of DCAA.

[46] In their study investigating the adverse male reproductive effects of subchronic exposure to DCAA, Toth and colleagues administered DCAA to male Long-Evans rats with doses of 1, 31.25, 62.5, or 125 mg/kg/day by oral gavage for 10 weeks (Toth et al., 1992). A decrease in sperm counts and an effect on sperm morphology were seen at the two highest doses. In addition a decrease in sperm motility, and inhibited sperm production in the testes was seen at the highest dose. Fertility, determined by the number of viable implants on Day 14 of gestation after an overnight mating, was decreased in the highest dose group.

[47] Linder and colleagues studied the spermatogenic effects of bromoacetic acids by dosing male rats with MBAA and DBAA (Linder et al., 1994). No effect on sperm was seen in the rats dosed with MBAA. A decrease in serum testosterone was observed two days after dosing with 1250 mg/kg DBAA. In addition, sperm motion was affected on days 14 and 28 after dosing, as well as abnormal sperm head shape and flagellar degeneration. Other effects such as decreased sperm counts were also observed. The authors of the study concluded that DBAA is a testicular toxicant to the rat.

(Smith et al., 1989b[48]), fetal developmental defects or malformations (Smith et al., 1989b; Christ et al., 1996[49]), maternal deaths (Smith et al., 1989b; Christ et al., 1996), and fetal deaths (Christ et al., 1996). HANs were also found to cross the placental barrier (Jacob et al., 1998[50]).

These studies suggest that DBPs are capable of causing various adverse reproductive effects, supporting the biological plausibility criterion of the criteria for causation. However, there are certain limitations in the application of these study results to human exposure to DBPs in drinking water. The first limitation stems from the interspecies applicability of the studies. The studies were all done on non-human mammals with much smaller body weights and different metabolic pathways than humans. The second limitation stems from the extremely high doses used in these studies compared to environmental exposures from drinking water. However, the mammals used in these studies generally have a much greater body surface area-to-body weight ratio compared to humans, which can translate into higher rates of metabolism. These higher metabolic rates may result in an ability to tolerate higher doses of toxins per kilogram body weight than humans can. In addition, the homogeneity of the animal study populations can not adequately address the effect of exposure on sensitive sub-groups and individuals that are likely to occur in the very heterogeneous human population. The final limitation is the number of animals used in toxicological studies. While environmental compounds may be associated with outcome rates of one in thousands, toxicological studies generally use fewer than thirty animals. Taking these limitations into account, a summary of the doses used in these studies and their calculated human equivalents can be found in Table 1-2.

---

[48] Smith and colleagues carried out two studies investigating the reproductive effect of HANs. In the first study pregnant Long-Evans rats were dosed by gavage with TCAN at 0, 1, 7.5, 15, 35 or 55 mg/kg/day in tricaprylin oil[48] during days 6-18 of gestation (Smith et al., 1988). Maternal deaths occurred in 21% of the rats at the highest dose and one death at the next lower dose. Complete litter reabsorptions were observed in two thirds of the survivors at the highest dose. Fetal weight and viability decreased with increasing dose. Malformations of the cardiovascular and urogenital systems were observed at all doses with the frequency of malformations increasing with increasing dose. In the second study, pregnant Long-Evans rats were administered DCAN by gavage at doses of 1, 5, 15, 25, or 45 mg/kg/day in tricaprylin oil (caprylic or octanoic acid triglyceride) as during days 6-18 of gestation (Smith et al., 1989b). Maternal deaths occurred in 9% of the rats at the highest dose and full litter reabsorption occurred in 60% of the survivors. Soft tissue and urogenital malformations as well as skeletal malformations occurred in a dose-dependent manner and were statistically significant at the highest dose. Embryolethality increased with increasing dose and was statistically significant at the two highest doses.

[49] A study on pregnant Long-Evans rats investigated the developmental effects of TCAN (Christ et al., 1996). The rats were dosed at 15, 35, 55 or 75 mg/kg/day in corn oil or 15 mg/kg/day in tricaprylin during days 6-18 of gestation. Maternal deaths occurred in one quarter of the rats at the highest dose and the percent of non-live implants increased at the two highest doses in corn oil. The dose-response curve for fetal effects, including cardiovascular malformations, shifted to the right (i.e. reduced toxic effects) when comparing doses administered in corn oil to those administered in tricaprylin. The study authors suggest that TCAN may interact with tricaprylin to produce a greater effect on fetal cardiovascular malformations.

[50] Jacob and colleagues undertook a study of the distribution of radio-labeled chloroacetonitrile in pregnant mice (Jacob et al., 1998). Mice were dosed intravenously at a dose equivalent to 77 mg/kg and processed for radiography at 1, 8, and 24 hours after dosing. Rapid uptake of radioactivity was seen in major maternal organs (including the liver, lungs, urinary bladder, gastrointestinal mucosa, cerebellum, and uterine laminar fluid), as well as in the liver and brain tissue of the fetuses. The study authors concluded that chloroacetonitrile or its metabolites are capable of crossing the placental barrier. This has important implications for the transport of HANs in the bodies of pregnant women.

Table 1-2. Doses from DBP toxicology studies and corresponding adult human doses

Study	Compound(s)	Study dose	Corresponding human dose[a,b]	Corresponding water ingested[c]
Thompson et al., 1974	TCM	2) Rats: 0, 20, 30, 126 mg/kg/day 2) Rabbits: 0, 20, 35, 50 mg/kg/day	2) 0, 1300, 1950, 8200 mg/day 2) 0, 1300, 2300, 3300 mg/day	1) 0, 13000, 19500, 82000 L/day 0, 13000, 23000, 33000 L/day
Ruddick et al., 1983	2) TCM 2) TBM, BDCM, CDBM	2) 100, 200, 400 mg/kg/day 2) 50, 100, 200 mg/kg/day	2) 6500, 13000, 26000 mg/day 2) 3300, 6500, 13000 mg/day	1) 65000, 130000, 260000 L/day 2) 33000, 65000, 130000 L/day
Narotsky et al., 1997	BDCM	0, 25, 50, 75 mg/kg/day	0, 1800, 3500, 5300 mg/day	0, 18000, 35000, 53000 L/day
Schwetz et al., 1974	TCM (inhalation)	30, 100, 300 ppm$_v$	380, 1300, 3800 mg/day	3800, 13000, 38000 L/day
Murray et al., 1979	TCM (inhalation)	0, 100 ppm$_v$	0, 1300 mg/day	0, 13000 L/day
Land et al., 1981	TCM (inhalation)	2) 0.08% (vol.) = 800 ppm$_v$ 2) 0.04% (vol.) = 400 ppm$_v$	2) 10000 mg/day 2) 5000 mg/day	1) 100000 L/day 2) 50000 L/day
Smith et al., 1989a	TCAA	1, 330, 800, 1200, 1800 mg/kg/day	65, 21000, 52000, 78000, 120000 mg/day	650, 210000, 520000, 780000, 1200000 L/day
Smith et al., 1992	DCAA	1) 0, 900, 1400, 1900, 2400 mg/kg/day 2) 1, 14, 140, 400 mg/kg/day	1) 0, 63000, 98000, 130000 mg/day 2) 70, 980, 9800, 28000 mg/day	1) 0, 630000, 980000, 1300000 L/day 2) 700, 9800, 98000, 280000 L/day
Epstein et al., 1992	DCAA	1900, 2400, 3500 mg/kg/day	130000, 170000, 250000 mg/day	1300000, 1700000, 2500000 L/day
Cicmanec et al., 1991	DCAA	0, 12.5, 39.5, 72 mg/kg/day	0, 880, 2800, 5000 mg/day	0, 8800, 28000, 50000 L/day
Toth et al, 1992	DCAA	1, 31.25, 62.5, 125 mg/kg/day	70, 2200, 4400, 8800 mg/day	700, 22000, 44000, 88000 L/day
Linder et al., 1994	2) MBAA 2) DBAA	2) 0, 100 mg/kg 2) 1, 1250 mg/kg	2) 0, 7000 mg/day 2) 70, 88000 mg/day	1) 0, 70000 L/day 2) 700, 880000 L/day

Table 1-2, continued. Doses from DBP toxicology studies and corresponding adult human doses

Study	Compound(s)	Study dose	Corresponding human dose[a,b]	Corresponding water ingested[c]
Smith et al., 1988	TCAN	0, 1, 7.5, 15, 35, 55 mg/kg/day	0, 70, 530, 1100, 2500, 3900 mg/day	0, 700, 5300, 11000, 25000, 39000 L/day
Smith et al., 1989b	DCAN	1, 5, 15, 25, 45 mg/kg/day	70, 350, 1100, 1800, 3200 mg/day	700, 3500, 11000, 18000, 32000 L/day
Christ et al., 1996	TCAN	15, 35, 55, 75 mg/kg/day	1100, 2500, 3900, 5300 mg/day	11000, 25000, 39000, 53000 L/day
Jacob et al., 1998	Chloro-acetonitrile	77 mg/kg	5400 mg/day	54000 L/day

[a] For ingestion, assuming the average adult weighs 70 kg.

[b] For inhalation, daily dose was obtained from ppm_v dose by following calculation: daily dose = $(\text{ppm}_v)*(\text{mol. wt. of TCM}*(19.2\ \text{m}^3/\text{day})/(24.45\ \text{L})$, where 24.45 L is generated from the Ideal Gas Law (PV = nRT; Petrucci, 1989) and is the volume of one mole of gas at 25 °C and 1 atm, and 19.2 m^3/day is the average inhalation rate of an adult based on 8 hours of resting and 16 hours of light activity (U.S. EPA, 1997). Units of ppm_v broken down are $(\text{mg}/\text{m}^3)*(24.45\ \text{L}*\text{g}/\text{mol})$

[c] This column represents the volume of water the average adult human would be required to consume on a daily basis to ingest the doses used in the studies, assuming a high concentration of 100 µg/L of the compounds of interest in the water. Calculated as: [human dose (mg/day) * 1000µg/mg] / [100 µg/L]

Studies attempting to quantify THM exposures have estimated daily exposure from all routes for the average adult to range between 100 and 370 µg for chloroform (Wallace et al., 1984; Wallace, 1997). Assuming the average adult weighs approximately 70 kg, this translates into 1.4 to 5.3 µg/kg/day, much lower than the doses applied in the toxicology studies. In contrast, a reproductive toxicology study on Fischer rats comparing bottled water to tap water collected from homes in Santa Clara County found no significant differences in adverse maternal or fetal effects in those rats fed tap water compared to the rats fed bottled water (Keen et al., 1992). Therefore, while the toxicology studies support the biological plausibility for several of the DBPs being causal agents for adverse reproductive effects, they do so only at doses unlikely to occur from chlorinated drinking water.

Exposure pathways

Exposure to an environmental chemical is a complicated issue and the language used to describe exposure is inconsistent. It is important to establish what is meant by "exposure" with respect to the environmental epidemiology studies discussed in this work. To this end, it is necessary to have a clear understanding of the mechanisms and semantics inherent in exposure assessment in order to avoid confusion.

For the purposes of exposure assessment, it is helpful to think of the human body as having a hypothetical outer boundary separating the inside of the body from the outside world (Paustenbach, 2000; U.S. EPA, 1992b). The process of a chemical entering a human body can be broken down into two parts: *exposure*, or the condition of the chemical coming into contact with the outer boundary of the human body; and *entry*, when the chemical actually crosses the boundary into the human body. Entry is itself broken down into *intake* and *uptake*. Intake describes the physical action of moving the chemical through an opening in the outer boundary layer, for example ingestion through the mouth or inhalation through the mouth or nose. Uptake describes the absorption of the chemical through the skin (via the dermal route) or across other barriers such as the gastro-intestinal lining in ingestion or the membranes in the respiratory tract in inhalation. Intake and uptake mechanisms are often grouped together and termed "uptake" (Paustenbach, 2000; U.S. EPA, 1992b). However, another concept in the exposure equation is that of dose which incorporates the concepts of intake and uptake. Dose is a measure of the amount of the chemical that crosses the boundary of the body and which may be absorbed or deposited in the body and ultimately over time reach the target site where the health effect is expected to occur (Hrudey et al., 1996).

The definition of dose as a general concept is very broad, so various types of dose have been defined: *external dose*[51] which is broken down into several subgroups including *applied dose*[52], *administered or potential dose*[53], *and exposure dose*[54]; *internal dose*[55]; *and biologically effective dose*[56]. Epidemiology studies that investigate both water concentrations and ingestion volume, along with other water use activities are ideally endeavoring to estimate a potential dose.

---

[51] The amount of a chemical at a site that is available for absorption and is the type of exposure least able to be used in the investigation of a dose-response relationship (Hrudey et al., 1996).

[52] The amount of a chemical that is available at the absorption barrier (Paustenbach, 2000; Hrudey et al., 1996; U.S. EPA, 1992b). Applied dose is most useful if it can be linked to internal dose in some way. In terms of the epidemiology studies of interest in this paper, applied dose corresponds to the concentration of DBP species in the drinking water of study subjects.

[53] The amount of a chemical that is actually ingested, inhaled or applied on the skin (Paustenbach, 2000; Hrudey et al., 1996; U.S. EPA, 1992b). It is the potential amount that could be absorbed IF the chemical is 100% bioavailable. However, the use of the administered or potential dose does not imply that 100% bioavailability or absorption is assumed (U.S. EPA, 1992b)

[54] The amount of the chemical in the environmental media to which an individual or a population is exposed over time (Hrudey et al., 1996).

[55] The amount of the chemical absorbed into the systemic circulation and available for interactions with the target organs (Paustenbach, 2000; Hrudey et al., 1996; U.S. EPA, 1992b).

[56] The amount of the chemical that actually reaches the target site and is therefore the most useful when assessing a dose-response relationship (Paustenbach, 2000; Hrudey et al., 1996). The biologically effective dose is very difficult to estimate, requiring either biological monitoring or mathematical modeling. Measures of the biologically effective dose generally cannot be accounted for by the exposure because the measured chemicals at the target site may not be specific to the unique exposure source or route of interest (Paustenbach, 2000; Hrudey et al., 1996).

There are three general methods used to quantify exposure: the *direct method*[57], the *reconstruction or biomonitoring* approach[58], and the *predictive or exposure scenario* approach[59] The predictive approach is the most common method of exposure assessment and is the method used in all of the reproductive epidemiology studies of interest here.

Most of the epidemiology studies of interest here have attempted to estimate the applied dose or the potential dose in what they termed their exposure assessment. For the balance of this paper, the words "exposure" and "exposure assessment" will be used to describe these estimations of dose employed by the epidemiology studies. To develop an accurate assessment of exposure in terms of a predictive approach to measuring applied or potential dose, it is necessary to investigate the three major routes of exposure to DBPs in chlorinated drinking water. These exposure routes are *ingestion*, *inhalation*, and *dermal contact*.

Ingestion exposure

Ingestion exposure occurs when food, beverages, or another carrier containing a chemical of interest are consumed. The chemical and the carrier pass through the esophagus into the stomach, small intestine, and colon before being excreted. During the passage through the gastro-intestinal (GI) tract, the carrier is broken down, and nutrients and the chemical of interest may be absorbed. The first destination for these absorbed compounds is the liver, which is served by circulation directly from the gastrointestinal tract, called the portal circulation. The liver is the first line of metabolism. Many chemical contaminants in food and water will be metabolised in the liver in what is called the "first pass effect" because the portal circulation provides the "first pass" to the liver before entering the systemic circulation. An illustrative case was determined by Weisel and Jo (Weisel and Jo, 1996) who found in human studies that an ingested dose of TCM could not be measured, suggesting that it was completely metabolized before it entered the systemic circulation. As a result of this first pass effect, many substances entering the body through the GI tract never enter the systemic circulation to be carried to their target site and may not appear to be bioavailable via this route of exposure. However, the metabolites of these compounds may enter the systemic circulation. This has implications for the determination of dose-response relationships, since the metabolic route is important in determining whether a chemical will reach its target site. Another important fact to consider is whether the parent chemical or its metabolite is the causal agent. The first pass effect is an efficient process for metabolising many parent chemicals and as a result may effectively deliver the causal agents, if those causal agents are the metabolites, to the systemic circulation and subsequently to the target sites.

---

[57] A measurement of the exposure concentration and duration of contact at the point of contact while the exposure is taking place.

[58] Which estimates past exposure through the use of internal indicators such as biomarkers, body burden, and excretion levels.

[59] Which sets up a hypothetical scenario in which information on concentrations of a chemical in one or more environmental phases is determined and combined with population characteristics, duration of contact, and activities of individuals or populations that are pertinent to the exposure (Paustenbach, 2000; Hrudey et al., 1996; U.S. EPA, 1992b).

Aside from the issue of bioavailability, exposure through ingestion depends on both the concentration of the chemicals in the carrier substance (food, beverages, soil, etc.) and the amount of the carrier substance ingested. Factors associated with concentrations of DBPs in water have already been discussed. However, there is substantial inter- and intra-individual variability in the amount of water or beverages made with water ingested in a population (Shimokura et al., 1998; Froese et al., Submitted; Bader et al., Submitted). The amount of water consumed by individuals in a study can be determined most accurately using daily journals[60] or, less accurately, using national data on consumption (U.S. EPA, 1997).

Inhalation exposure

Inhalation exposure occurs in the respiratory tract, which is divided into three regions. These regions are the nasopharyngeal region[61]; the tracheobronchial region[62], and the pulmonary region[63]. The pulmonary region is the gateway to the circulatory system where the air-blood barrier is very accessible due to the large surface area. This is where most of the absorption from inhalation exposure occurs. The physical form of the exposure chemicals entering the respiratory tract is important in determining bioavailability. If the chemicals are in a physical form that precludes their arriving at the pulmonary region, they will not be absorbed into the body as readily. For example, particles greater than 10μm in diameter tend to be captured in the upper respiratory tract or nasopharyngeal region and are then expelled from the respiratory tract and ingested. Chemicals attached to these particles would not be bioavailable through the inhalation route. Inhalation exposure to chemicals in drinking water will occur mainly through volatile compounds in the vapor phase and dissolved compounds in aerosols of variable particle size.

The release of volatile chemicals from water depends on the concentration gradient between the water and the air around it, the temperature, and the Henry's Law constant for the compound (Mercer, 1999)[64].

---

[60] Daily journals focus on the amount of direct consumption of hot or cold tap water, beverages made with hot or cold tap water, or food prepared with hot or cold tap water. Questions on relative amounts of tap water and bottled water should be included. Other sources of DBPs that are difficult to quantify are purchased, prepared beverages such as soft drinks.

[61] Comprising the nose, nasopharynx and the larynx;

[62] Comprising the trachea, bronchi and bronchioles

[63] Comprising the respiratory bronchioles, alveolar ducts, alveolar sacs and alveoli

[64] The Henry's Law constant (H) is also known as the partition coefficient between the atmosphere and water and is calculated as: H = (the concentration of the compound in the atmosphere at equilibrium)/ (the concentration of the compound in water at equilibrium). The Henry's Law constant is a ratio of the equilibrium concentrations of the compound of interest in air and in water. The Henry's Law constant gives a measure of whether a compound is more likely to be found in air or in water at equilibrium or in other words, the volatility of the compound. The greater the Henry's Law constant for a compound, the more volatile a compound is considered to be.

Exposure to chemicals in aerosols depends on the amount of aerosols produced, the size distribution of aerosols, the growth or shrinkage of aerosols, and aerosol transport and removal in the home. It is assumed that vapors or chemicals in aerosols that enter the pulmonary region will be absorbed, unless the chemicals are very insoluble (Paustenbach, 2000).

Chemicals are volatilized from water and aerosols are produced mainly when water is agitated in some way. The major source of volatile gases and aerosol formation in the home is the shower (Weisel et al., 1999b; Mercer, 1999). Other sources include bathing, washing clothes and dishes by hand or by machine, washing pets, cars and children, cooking, humidifiers, and flushing toilets.

Any inhalation exposure must account for the inhalation rate of the person being exposed, which can vary depending on such factors as gender, physical fitness, general health, body size, and activity patterns. Other factors that have a bearing on inhalation exposure include the concentration of DBPs in the water, the temperature of the water, the duration of the exposure, characteristics of the home[65], as well as water use by other members of the household (Weisel et al., 1999b).

Dermal exposure

Uptake of compounds through the skin is a complicated process. The skin is comprised of several layers. The outermost layer is the stratum corneum, a layer of dead, impermeable cells surrounded by a more permeable lipid bilayer. The stratum corneum is lipophilic. The next layer is the viable epidermis, made up of nucleated cells. The viable epidermis is hydrophilic. The final layer is the capillary network, which lies under the viable epidermis and leads diffused chemicals into the circulatory system. In addition, skin structures such as sweat glands, hair follicles and sebaceous ducts also have an effect on uptake through the skin by providing routes of penetration through the skin layers (Bunge and McDougal, 1999).

Generally, the rate of uptake through the skin is described by Fick's Law of Diffusion, which states that the rate of diffusion across a membrane is dependent on the diffusion coefficient of the membrane and the concentration gradient of the compound across the membrane[66].

---

[65] Including air exchange rates between rooms and between the inside of the house and the outdoors, the presence or absence of ventilation in the bathroom, and the location and frequency of use of sinks, washing machines, and dishwashers.

[66] An equation for Fick's Law modified for dermal uptake is as follows (Paustenbach, 2000): $J = dQ/dt = DK\nabla C/e \approx K_pC$ Where: J & dQ/dt = the chemical flux or rate of chemical absorbed (mg/cm^2h); D = diffusivity in the stratum corneum (cm^2/h); K = the stratum corneum/water (or other vehicle) partition coefficient (unitless); ∇C = the concentration gradient or difference between the concentration above and below the stratum corneum (mg/cm^3); e = the thickness of the stratum corneum (cm); K_p = permeability coefficient (cm/h); C = the applied chemical concentration (mg/cm^3).

Explaining Fick's Law in terms of dermal exposure, the diffusion coefficient is replaced by the permeability coefficient, which is a measure of a chemical's ability to penetrate across the skin, and the skin is assumed to be a homogenous membrane with the properties of the stratum corneum. In addition, models using Fick's Law assume that the chemical concentration outside the skin does not change during the time of the exposure, and there is no increase in the chemical concentration in the circulatory system. Accordingly, the rate of dermal uptake depends on the concentration of the chemical in the exposure vehicle, in our case the concentration of DBPs in water near the skin, and the permeability coefficient of the chemical into the stratum corneum. While Fick's Law should be regarded as only a rough estimate of the process of dermal exposure, it does provide a useful model for understanding the dermal uptake process.

The factors affecting water concentrations of DBPs have been discussed previously. The permeability coefficient rates further investigation. The permeability coefficient is a time- and concentration-independent estimation of dermal absorption of a chemical over the thickness of the skin and essentially is a partition coefficient between the exposure vehicle (i.e. water) and the skin. The U.S. EPA (U.S. EPA, 1992a) employed the octanol-water partition coefficient to estimate the permeability coefficient for approximately 200, mostly non-ionized, chemicals with the following equation:

$$\mathrm{Log}_{10} K_p = -2.72 + 0.71 \log_{10} K_{ow} - 0.0061\ MW$$

Where:
K_p = the permeability coefficient
K_{ow} = the octanol-water partition coefficient
MW = the molecular weight of the compound

Therefore, knowledge of a chemical's octanol-water partition coefficient aids in the estimation of the rate of dermal uptake of that chemical.

While the stratum corneum is generally considered the rate-limiting barrier to dermal uptake, there are other factors that influence the rate of dermal uptake. The ionic state of a chemical has a significant effect on the rate of dermal absorption since unionized species penetrate better than ionized species[67]. External temperature also has an effect on the rate of uptake of chemicals from water[68].

---

[67] However, structures in the skin such as sweat glands, sebaceous ducts and hair follicles can provide a route of entry for ionized species. The concentrations of these structures varies over different regions of the skin so the region of exposure can affect the uptake rate of chemicals depending on the ionization state of the chemicals.

[68] A study on dermal exposure to chloroform in chlorinated water (Gordon et al., 1998) found that an increase in water temperature resulted in increased absorption of chloroform as measured by the amount of chloroform in exhaled breath. They found that approximately 30 times more chloroform was exhaled by their study subjects between the highest and lowest temperatures investigated. The authors suggest that a likely explanation is that blood flow to the skin decreases

Other factors that may contribute to variations in the rate of dermal uptake include regional variations in thickness of the skin, hydration of the skin, and the use of soaps and surfactants. However, the effects of these factors are difficult to quantify.

Dermal exposure to DBPs in drinking water occurs via the same water use activities as inhalation exposure. Vapors and aerosols contribute to dermal exposure; however, the primary vehicle of exposure is liquid water and the primary exposure activities are showering and bathing.

Relative contributions from the various routes of exposure

Ingestion has long been considered the primary route of exposure to chemicals in drinking water. Most epidemiology studies that have tried to quantify exposure values have concentrated on ingestion exposure. However, an examination of the Henry's Law constants and octanol-water partition coefficients of two of the DBP compounds provides an interesting insight into the possible contribution of other routes of uptake into the body. The Henry's Law constants for TCM and TCAA are 430 Pa m^3/mol (Mackay et al., 1993) and 0.002 Pa m^3/mol [69], respectively. This suggests that TCM, being quite volatile, will likely volatilize from drinking water and as a result may enter the body via the inhalation route. TCAA, however, is likely to remain in the water phase. This 100,000-fold difference in volatility underscores how differences in chemical-physical properties can help to predict differing behaviours (and exposure routes) when comparing chemical compounds. In contrast, the log K_{ow}'s for these two compounds are quite similar: TCM log K_{ow} =1.97 (Mackay et al., 1993), TCAA log K_{ow} =1.33 (Mackay et al., 1995; Howard, 1997). As mentioned previously, a knowledge of the K_{ow} of a compound can aid in the estimation of the rate of dermal uptake of the chemical. Generally, the greater the K_{ow} the more quickly a chemical may be taken up dermally. TCM and TCAA have K_{ow}'s within an order of magnitude of each other and may behave very similarly (all other factor being equal) in terms of dermal uptake. In this case, the log K_{ow} values suggest that dermal uptake may be a notable exposure route for both compounds.

Several studies have suggested that dermal and inhalation exposure may contribute a substantial proportion of the uptake of certain DBPs from drinking water, and that ingestion exposure as calculated in epidemiology studies does not provide the whole picture for these DBPs. One study (Wallace et al., 1986) found that measuring the concentrations of volatile compounds in exhaled breath was a feasible method for determining environmental and occupation exposures to volatile compounds[70]. Recent studies re-confirmed the use of THM concentrations in exhaled breath as a biomarker of exposure for occupational (Fantuzzi et al.,

---

at cooler water temperatures in order to keep the core body temperature constant, resulting in less opportunity for the chloroform to enter the systemic circulation.

[69] Estimated from bond strength using a computer program.

[70] This method has been applied to exposures to volatile DBPs, in particular chloroform and the other THMs.

2001) and shower (Weisel et al., 1999a) exposure to THMs by correlating THM concentrations in exhaled breath with THM concentrations in ambient air or water.

Several studies have used this relationship to investigate the relative exposure contributions of the different routes of exposure. In one study (Jo et al., 1990) the relative contributions of inhalation and dermal exposure in a showering situation were investigated. TCM concentrations in exhaled breath were measured after a shower in which the subject was exposed both dermally and via inhalation (normal showering conditions) and again after a shower in which the subject wore waterproof clothing[71] and was only exposed via inhalation. They found that the relative contributions to exposure were equal from the dermal and inhalation routes. A similar study (Levesque et al., 1994) was carried out in a swimming pool. In this study, the TCM concentrations in exhaled breath after normal exposure in a swimming pool were compared to TCM concentrations in exhaled breath after only dermal exposure[72]. It was found that the dermal route contributed approximately 24% of the exposure in a swimming pool situation.

Weisel and Jo (Weisel and Jo, 1996) used TCM concentrations in exhaled breath to investigate the relative contributions of inhalation and dermal exposures compared to ingestion exposure. They found that the ingested dose yielded no TCM in exhaled breath and speculated that TCM was completely metabolized before it entered the systemic circulation due to the first pass effect, and was therefore not available for excretion through the breath[73]. They also estimated that dermal and inhalation routes contributed as much or more of the total exposure to TCM in water.

A study in Taiwan (Lin and Hoang, 2000) calculated the relative contributions of ingestion and inhalation exposure from models based on daily water use activities and estimated THM concentrations. These researchers modeled the amount of inhalation exposure in an average day from three water use activities: showering, pre-and post cooking activities, and cooking processes. They discovered that total inhalation exposure, summed from the three water use activities, was comparable to ingestion exposure (30.7 μg/day vs. 47.9 μg/day, respectively).

---

[71] In order to block dermal exposure

[72] Dermal exposure was isolated by having the subjects breathe through scuba tanks containing TCM-free air while swimming in the pool in normal bathing attire. This also ensured that no pool water was swallowed during the exposure. Normal exposure was defined by having the subjects exhibit normal swimming behaviour, breathing the air in the pool area as well as wearing normal swimming attire and possibly swallowing water.

[73] This is particularly significant in terms of the biologically effective dose. It seems evident from the use of TCM in exhaled breath as a biomarker that inhalation and dermal exposure leads to TCM being available to the systemic circulation where it is then available to many different target sites. On the other hand, uptake through ingestion leads to rapid metabolism of the TCM. Therefore, if the target site for TCM is the liver and TCM is the agent of interest, then ingestion exposure is the primary route to determine biologically effective dose. However, if other organs are the target sites, as it would seem in the investigation of reproductive effects, then dermal and inhalation routes should be investigated. In contrast, if the metabolites of TCM are the agents of interest, then inhalation and dermal exposure would not seem to provide as effective a dose as ingestion exposure, which would ensure metabolism of the TCM. Therefore, it is necessary to determine whether the causal agent is the parent compound or its metabolites before the route delivering a biologically effecting dose can be determined

A recent study in the United States (Lynberg et al., 2001) investigated an association between THM concentrations in blood and THM intake from showering and ingestion. Blood samples were collected from study subjects before and after showering. The two groups of subjects lived in locations with very different water quality in terms of THM speciation and concentrations. The researchers found that THM speciation and concentrations in the blood mirror the THM speciation and concentrations in the water supply at both background (pre-shower) and after shower levels.

While the majority of studies investigated THM exposure by looking at THM concentrations in exhaled breath, another study (Kim and Weisel, 1998) investigated haloacetic acids (HAAs) exposure via ingestion by looking at HAA concentrations in urine. Subjects recorded their water intake and other potential exposures in a 48-hour recall questionnaire. First morning urine samples as well as tap water samples were obtained and analyzed. A significant correlation was found between ingestion exposure to TCAA and urinary excretion rates of TCAA, suggesting that TCAA could be used as a biomarker of exposure.

Two recent studies, both by the Environmental Health group at the University of Alberta, have been carried out to investigate further the relationship between TCAA ingestion exposure and TCAA urinary excretion rates. The first of these studies was a pilot study carried out in Adelaide, Australia (Froese et al., 2002, In Press). Ten volunteers were recruited from the water quality research laboratory in Adelaide. The goals of the study were to investigate inter- and intra-individual variability in TCAA ingestion and urinary excretion. TCAA concentrations in tap water at the water quality laboratory as well as the volunteers' homes were analyzed daily, and the volunteers kept detailed consumption[74] and exposure diaries over the five-week study period. First morning urine samples were also analysed daily. The results showed substantial intra- and inter-individual variability in both TCAA ingestion and excretion rates from day-to-day. A substantial portion of the TCAA ingestion variability was the result of variability in TCAA concentrations in the source water from day-to-day, as well as variability in the volume of water consumed. Part way through the study, the volunteers were asked to consume and prepare beverages with only TCAA-free bottled water to determine a possible elimination half-life[75] for TCAA. Data from only three of the volunteers were used for this part of the study; however, elimination half-lives were found to be 3.7 days, 2.3 days, and 2.9 days[76]. This study confirmed that TCAA is readily detectable in urine with the analytical methods used in this study and TCAA is stable enough in urine to allow for reasonable turn-around time for analysis, characteristics required in a biomarker. It was determined that future research on TCAA excretion as a biomarker for ingestion should control the TCAA ingestion rates of the study volunteers.

---

[74] Consumption included cold and hot tap water, including beverages prepared with boiled tap water, and commercially prepared beverages. Other exposures such as showering/bathing, washing dishes/ cars, swimming, etc. were detailed in the exposure diaries, as well as exposure to dry-cleaning solvents.

[75] The elimination half-life ($t_{1/2}$)is the time required for the concentration of a chemical in the body to decrease to half of its original value. The half-life can be calculated from the elimination rate

To this end, a second study was carried out in Edmonton, Alberta, Canada, also by the Environmental Health group at the University of Alberta (Bader et al., Submitted). Five volunteers were recruited to consume solely Winnipeg treated drinking water for two weeks and TCAA-free bottled water for an additional two weeks[77]. First morning urine samples were collected daily and detailed consumption diaries were kept by each volunteer. Elimination half-lives were calculated using creatinine-normalized[78] first morning urine excretion for each volunteer and found to be 2.1, 2.3, 2.5, 5.0, and 6.2 days[79]. It was found that creatinine-normalized TCAA excretion may provide a more accurate method of determining TCAA elimination than TCAA excretion that is not creatinine-normalized.

Biomarkers

As the previous sections have shown, there are many factors contributing to sources of variation and error in the calculation of exposure to disinfection by-products in drinking water. Therefore, it would seem that biomarkers of exposure are a promising solution to the difficulties inherent in the determination of exposure assessment by more routine methods. Indeed, biomarkers of exposure have been suggested as a "gold standard" in determining exposure to disinfection by-products in drinking water (Swan and Waller, 1998). The question arises, then: why have no epidemiology studies used biomarkers? The answer to this question lies in the list of characteristics necessary for the ideal biomarker of exposure. Before examining these characteristics, it is necessary to understand the different types of biomarkers.

Biomarkers are classified as one of three types (U.S. National Research Council, 1989): *biomarkers of exposure*[80], which are subclassified as *biomarkers of*

---

constant (k_{el}), which in turn is ascertained by plotting the log of the concentration in the body against time. The elimination half-life is determined using the equation: $t_{1/2} = 0.693/k_{el}$.

[76] R^2 values of 0.45, 0.74 and 0.70, respectively

[77] Winnipeg water was imported as the TCAA source. Because TCAA levels in Edmonton water are very low compared to TCAA levels in Winnipeg water (Rizak et al., 2000), it was thought that the Winnipeg water would provide better response in terms of urinary TCAA excretion.

[78] Creatinine is a waste product of normal muscle metabolism. It is produced and eliminated in the urine at a constant rate. The amount of creatinine produced and eliminated is not affected by exercise, fluid consumption, temperature, or diet. As such, it can be used to account for changes in urine volume.

[79] R^2 values of 0.83, 0.94, 0.88, 0.80, and 0.76, respectively

[80] The identification of extraneous compounds or contaminants within the biological system, or products of the interaction between the extraneous compounds and components of the host

internal dose[81] and *biomarkers of biologically effective dose*[82]; *biomarkers of effect*[83]; and *biomarkers of susceptibility*[84].

The biomarkers of interest in these reproductive epidemiology studies are biomarkers of exposure. No biomarkers of effect or susceptibility have been proposed to date. There are certain characteristics, which determine the feasibility of a biomarker of exposure. The first of these characteristics is the time to appearance, or the time it takes after exposure for the biomarker to be measurable in the body. The time to appearance depends on the part of the body being measured[85], as well as the metabolic pathway of the substance in the body (Wilcosky, 1990). It is preferable to choose a biomarker that will be measurable soon after the exposure occurs.

Linked to the time of appearance is the persistence of the biomarker in the body. The length of time the biomarker remains in the body determines whether the biomarker can be used to measure past, recent, or ongoing exposures. Most biomarkers will eventually decrease in concentration in the body after exposure ceases as the biomarker is metabolized and excreted. If on-going, cumulative

---

biological system (U.S. National Research Council, 1987). Unlike traditional methods of exposure determination which rely on an estimation of external dose by modeling or ambient monitoring, biomarkers of exposure give a measure of internal dose, that is, the amount of the extraneous compound or contaminant of concern that has been absorbed into the system.

[81] Which measure the concentration of the contaminant or its metabolite in cells, tissues or body fluids. From this information, a good start can be made towards determining an exposure-response relationship as the exposure is quantified more accurately.

[82] Which indicates that the contaminant of concern or its metabolite has reached a target site in the system. However, if the metabolites of the contaminant of concern must be analysed rather than the parent compound, then the biomarker may not be related as directly to the exposure (U.S. National Research Council, 1991).

[83] These are the next step removed from the actual exposure, but conversely, are most useful for determining health effects. Biomarkers of effect are identified as "any change that is qualitatively or quantitatively predictive of health impairment or potential impairment resulting from exposure" (U.S. National Research Council, 1987). This can include the induction or suppression of certain chemicals or actual physiological changes in the body and is often a reversible subclinical manifestation of disease. Biomarkers of effect are very useful in risk assessment if they can be confidently related to the exposure suspected to be responsible for an effect. Unfortunately, it is often difficult to relate the subclinical symptoms of disease with the environmental exposure, particularly if there is a significant time lag between exposure and effect.

[84] Individual response to an exposure depends in part on the inherited and acquired characteristics of the individual, which are termed biomarkers of susceptibility, the third class of biomarkers. Biomarkers of susceptibility are measures of how sensitive an individual is to the challenges posed by a contaminant of concern. These can include genetic characteristics as well as acquired characteristics such as a pre-existing disease, physiological change, or medication that can render the individual less resistant to the effects of exposure to the contaminant. An example of a genetic characteristic that serves as a biomarker of susceptibility is the sickle cell trait which predisposes an individual to anemia and altitude sickness as well as sensitivity to carbon monoxide, cyanide and aromatic amino and nitro compounds, but which provides some protection from malaria. Acquired biomarkers of susceptibility include antigen-specific antibodies that are developed by the body when a prior exposure to a chemical has induced an immunological response that sensitizes the individual to subsequent exposures (World Health Organization, 1993).

[85] i.e. fluids such as blood, urine or saliva, or tissues such as organs or muscles

exposures are to be measured, it is desirable to pick a biomarker with an elimination half-life of days rather than of hours (Wilcosky, 1990). In addition, the biomarker should be stable enough in the medium sampled that it can be measured.

The biomarker must also be sensitive to the exposure of interest. In other words, the biomarker must be detectable in exposed populations and undetectable or present in low levels in unexposed populations. The biomarker must also be sensitive enough to be able to detect exposure differences. The biomarker must be specific to the exposure of interest; a good biomarker of exposure should be able to identify and measure only the exposure of interest. It must be biologically relevant to the exposure of interest and the pharmacokinetic behaviours of the biomarker[86] should be well-characterized (U.S. National Research Council, 1989). In addition, biomarkers of exposure must be feasible to use. In other words, sample collection must not be too invasive, and measurement of the biomarker must not be economically and technologically prohibitive.

As mentioned previously, THMs in breath and blood and HAAs in urine have been put forward as possible biomarkers of exposure to DBPs in chlorinated drinking water. THMs have a very short half-life in the body, usually on the order of hours. This means that THMs cannot be used for cumulative exposures over long periods of time, although the results of a recent study (Lynberg et al., 2001) suggest that there are background levels of THM concentrations in the blood that may be promising for use as biomarkers. TCAA in urine has been found experimentally to have an elimination half-life of between 2.1 and 6.2 days (Froese et al., 2002, In Press; Bader et al., Submitted). Three studies have investigated a correlation between TCAA in urine and TCAA exposure from drinking water (Weisel et al., 1999a; Froese et al., 2002, In Press; Bader et al., Submitted). TCAA is presently the most promising candidate for a biomarker of exposure to DBPs in drinking water. However, TCAA is not an ideal biomarker of exposure[87].

The validation of appropriate and feasible biomarkers of exposure to DBPs in chlorinated drinking water continues. However, the questions about the potential link between adverse reproductive effects and exposure to DBPs in chlorinated drinking water remain and epidemiology studies continue. Therefore, exposure assessment by more mundane means remains an important element in epidemiology studies. Swan and Waller (Swan and Waller, 1998) have this to say about exposure assessment:

---

[86] Absorption, distribution, metabolism, and excretion

[87] Collection and analysis of urine samples is time consuming and expensive. In addition, TCAA is a known metabolite of both 1,1,2-trichloroethylene (TCE) and tetrachloroethylene exposure and its presence in urine is therefore not exclusive to TCAA exposure from drinking water (Crebelli and Carere, 1989). Exposure of the general population to tetrachloroethylene occurs mainly because of exposure to dry cleaning chemicals. Exposure to TCE occurs mainly from its use as a degreasing agent in garages, industry, and some domestic cleaning products. These exposures can be accounted for in exposure-activity questionnaires.

> "Ideally, biomarkers of exposure to disinfection by-products would be used, but these have yet to be developed. Short of this gold standard, in future studies, consumption of tapwater and drinks made from tapwater at home, work, and elsewhere should be ascertained, and other water uses should be quantified. Factors modifying exposure, such as filtration, boiling, or refrigeration, should be ascertained. Estimates of individual water constituents must be obtained at the appropriate time and place, ideally in prospective studies with home sampling. Finally, exposure to a variety of disinfection by-products must be examined rather than restricting attention to TTHM."

Accordingly, exposure assessment is an important element in epidemiology studies and will be for some time to come. The questions posed at the beginning of this section are particularly applicable in determining the state of exposure assessment in epidemiology studies to date and in developing pathways for future research. To recap, the questions are:

-How well have reproductive epidemiology studies to date characterized individual exposure to possible causal agents in treated drinking water and have there been useful improvements in exposure assessment over time?

-Can improved exposure assessment clarify a possible causal association between exposure to agents in treated drinking water and adverse reproductive effects?

-How can future epidemiology studies improve exposure assessment in the absence of biomarkers of exposure?

These questions will be addressed by the research presented in subsequent sections of this body of work.

References

Alvanja, M., I. Godstein and Susser M., 1978. A Case-control study of gastrointestinal and urinary tract cancer mortality and drinking water chlorination. in *Water Chlorination: Environmental Impact and Health Effects*. R. L. Jolley, H. Gorchev and D. H. Hamilton. Ann Arbor (MI), Ann Armor Science Publishers: 395-409.

Bader, Erin, Kenneth L. Froese and Steve E. Hrudey, Submitted. Determination of Elimination Half-Life for TCAA in a Controlled Human Exposure Study. .

Balko, Justin, Kenneth L. Froese and Steve E. Hrudey, 2001. *Is Drinking Coffee Worse Than We Think? Production of HAAs During Beverage Preparation*. Microbial/Disinfection By-Products Health Effects Symposium, Lisle, Illinois, American Water Works Association.

Beaglehole, R., R. Bonita and T. Kjellstrom, 1993. *Basic Epidemiology*. Geneva, World Health Organization.

Benoit, Frank M., Guy L. LeBel and Brian Jay, 2000. Evaluation of Techniques for the Removal of Chlorinated Disinfection By-Products form Drinking Water at the Point of Use. *Presented as a poster at An International Workshop on Exposure Assessment for Disinfection By-Products in Epidemiology Studies. Ottawa, Canada. May 7-10, 2000.*

Brenniman, G.R., J. Vasilomanolakis-Lagos, J. Amsel, N. Tsukasa and A.H. Wolff, 1980. Case-Control study of cancer deaths in Illinois communities served by chlorinated or non-chlorinated water. in *Water Chlorination: Environmental Impact and Health Effects*. R. L. Jolley, W. A. Brungs and R. B. Cumming. Ann Arbor (MI), Ann Arbor Science Publishers: 1043-1057.

Bunge, Annette L. and James N. McDougal, 1999. Dermal Uptake. in *Exposure to Contaminants in Drinking Water: Estimating Uptake through the Skin and by Inhalation*. S. S. Olin. Washington, D.C., CRC Press/ International Life Sciences Institute Press.

Cantor, K.P., R. Hoover, Harge P., Mason T.J., Silverman D.T. and Altman R., 1987. Bladder cancer, drinking water source and tap water consumption: a case-control study. *Journal of the National Cancer Institute* **79**: 1269-1279.

Christ, Suzanne A., Eleanor J. Read, Judy A. Stober and M. Kate Smith, 1996. Developmental Effects of Trichloroacetonitrile Administered in Corn Oil to Pregnant Long-Evans Rats. *Journal of Toxicology and Environmental Health* **47**: 233-247.

Cicmanec, J.L., L.W. Condie, G.R. Olson and Shin-Ru Wang, 1991. 90-Day Toxicity Study of Dichloroacetate in Dogs. *Fundamental and applied Toxicology* **17**: 376-389.

Crebelli, Riccardo and Angelo Carere, 1989. Genetic Toxicology of 1,1,2-trichloroethylene. *Mutation Research* **221**: 11-37.

Deane, Margaret, Shanna H. Swan, John A. Harris, David M. Epstein and Raymond R. Neutra, 1989. Adverse Pregnancy Outcomes in Relation to Water Contamination, Santa Clara County, California, 1980-1981. *American Journal of Epidemiology* **129**(5): 894-904.

Diehl, Alicia C., Gerald E. Speitel, James M. Symons, Stuart W. Krasner, Cordelia J. Hwang and Sylvia E. Barret, 2000. DBP Formation During Chloramination. *Journal of the American Water Works Association* **92**(6): 76-90.

Diehl, Alicia C., Gerard E. Jr. Speitel, James M. Sumons and Stuart W. Krasner, 1995. *Factors Affecting Disinfection By-Product Formation During Chloramination*. Water Quality Annual conference, Anaheim, California, American Water Works Association.

Epstein, D.L., G.A. Nolen, J.L. Randall, S.A. Christ, E.J. Read, J.A. Stoberet al., 1992. Cardiopathic Effect of dichloroacetate in the Fetal Long-Evans Rat. *Teratology* **46**: 225-235.

Fantuzzi, Guglielmina, Elena Righi, Guerrino Predieri, Giorgia Ceppelli, Fabriziomaria Gobba and Gabriella Aggazzotti, 2001. Occupational Exposure to Trihalomethanes in Indoor Swimming Pools. *The Science of the Total Environment* **264**: 257-265.

Freedman, D.M., K.P. Cantor, Lee N.L., L.S. Chen, H.H. Lei and C.E. Ruhl, 1997. Bladder cancer and drinking water: a population-based case-control study in Washington County, Maryland. *Cancer Causes and Control* **8**: 738-744.

Froese, Kenneth L., Martha I. Sinclair and Steve E. Hrudey, 2002, In Press. Trichloroacetic Acid as a Biomarker of Human Exposure to DBPs in Drinking Water - Adelaide Trial. *Environmental Health Perspectives* **110**.

Gordon, Sydney M., Lance A. Wallace, Patrick J. Callahan, Conald V. Kenny and Marielle C. Brinkman, 1998. Effect of Water Temperature on Dermal Exposure to Chloroform. *Environmental Health Perspectives* **106**(6): 337-345.

Gottleib, M.S., J.K. Carr and J.R. Clarkson, 1982. Drinking water and cancer in Louisiana: a retrospective mortality study. *American Journal of Epidemiology* **116**: 652-667.

Graves, Carol G., Genevieve M. Matanoski and Robert G. Tardiff, 2001. Weight of Evidence for an Association between Adverse Reproductive and Developmental Effects and Exposure to Disinfection By-products: A Critical Review. *Regulatory Toxicology and Pharmacology* **34**: 103-124.

Hennekens, Charles H. and Julie E. Buring, 1987. *Epidemiology in Medicine*. Boston, Little, Brown and company.

Hill, Austin Bradford, 1965. The Environment and Disease: Association or Causation? *Proceedings of the Royal Society of Medicine* **58**: 295-300.

Howard, Philip H., 1997. *Handbook of Environmental Fate and Exposure Data for Organic Chemicals*. Chelsea, Michigan, Lewis Publishers, Inc.

Hrudey, S.E., W. Chen and C. Rousseaux, 1996. *Bioavailability*. New York, CRC-Lewis Publishers.

Jacob, Sam, Abdel-Aziz H. Abdel-Aziz, Samia A. Shouman and Ahmed E. Ahmed, 1998. Effect of Glutathione Modulation on the Distribution and Transplacental Uptake of 2-{14C}-chloroacetonitrile (CAN) Quantitative Whole-Body Autoradiographic Study in Pregnant Mice. *Toxicology and Industrial Health* **14**(4): 533-546.

Jo, Wan K., Clifford P. Weisel and Paul J. Lioy, 1990. Routes of Chloroform Exposure and Body Burden from Showering with Chlorinated Tap Water. *Risk Analysis* **10**(4): 575-585.

Keen, Carl L., Janet Y. Uriu-Hare, Shanna H. Swan and Raymond R. Neutra, 1992. The Effects of Water Source on Reproductive Outcome in Fischer-344 Rats. *Epidemiology* **3**(2): 130-133.

Kim, Hekap and Clifford P. Weisel, 1998. Dermal Absorption of Dichloro-and Trichloroacetic Acids from Chlorinated Water. *Journal of Exposure Analysis and Environmental Epidemiology* **8**(4): 555-575.

King, Will D. and Loraine D. Marrett, 1996. Case-Control Study of Bladder Cancer and Chlorination By-Products in Treated Water (Ontario, Canada). *Cancer Causes and Control* **7**(6): 596-604.

Krasner, Stuart W., 1999a. *Chemistry and Occurrence of Disinfection By-Products*. The Safety of Water Disinfection: Balancing Chemical and Microbial Risks, Miami Beach, Florida, International Life Sciences Institute.

Krasner, Stuart W., 1999b. Chemistry of Disinfection By-Product Formation. in *Formation and control of Disinfection By-Products in Drinking Water*. P. C. Singer. Denver, Colorado, American Water Works Association: 424.

Krasner, Stuart W., Michael J. McGuire, Joseph G. Jacangelo, Nancy L. Patania, Kevin M. Reagan and E. Marco Aieta, 1989. The Occurrence of Disinfection By-Products in US Drinking Water. *Journal of the American Water Works Association* **81**(8): 41-53.

Lahl, Uwe, Mehmet Cetinkaya, Jurgen V. Duszeln, Bernd Gabel, Burkhard Stachel and Wolfram Thiemann, 1982. Health Risks for Infants Caused by Trihalomethane Generation During Chemical Disinfection of Feeding Utensils. *Ecology of Food and Nutrition* **12**: 7-17.

Land, Paul C., E.L. Owen and Harry W. Linde, 1981. Morphologic Changes in Mouse Spermatozoa after Exposure to Inhalational Anaesthetics during Early Spermatogenesis. *Anesthiology* **53**: 53-56.

Levesque, Benoit, Pierre Ayotte, Alain LeBlanc, Erid Dewailly, Denis Prud'Homme, Robert Lavoie et al., 1994. Evaluation of Dermal Respiratory Chloroform Exposure in Humans. *Environmental Health Perspectives* **102**(12): 1082-1087.

Lin, Tsair-Fuh and Shih-Wen Hoang, 2000. Inhalation Exposure to THMs from Drinking Water in South Taiwan. *The Science of the Total Environment* **246**: 41-49.

Linder, Ralph E., Gary R. Klinefelter, Lillian F. Strader, Juan D. Suarez and Cheryl J. Dyer, 1994. Acute Spermatogenic Effects of Bromoacetic Acids. *Fundamental and Applied Toxicology* **22**: 422-430.

Lynberg, Michele, J.R. Nuckols, Peter Langlois, David Ashley, Philip Singer, Pauline Mendola et al., 2001. Assessing Exposure to Disinfection By-products in Women of Reproductive Age Living in Corpus Christi, Texas, and Cobb County, Georgia: Descriptive Results and Methods. *Environmental Health Perspectives* **109**(6): 597-604.

Mackay, Donald, Wany Ying Shiu and Kuo Ching Ma, 1993. *Illustrated Handbook of Physical Chemical Properties and Environmental Fate for Organic Chemicals*. Boca Raton, Lewis Publishers.

Mackay, Donald, Wan-Ying Shiu and Kuo-Ching Ma, 1995. *Illustrated Handbook of Physical and chemical Properties and Environmental Fate for organic Chemicals*. New York, CRC Lewis Publishers.

McGeehin, Michael A. , John S. Reif, Judith C. Becher and Ellen J. Mangione, 1993. Case-Control Study of Bladder Cancer and Water Disinfection Methods in Colorado. *American Journal of Epidemiology* **138**(7): 492-501.

Mercer, Robert R., 1999. Respiratory Uptake. in *Exposure to Contaminants in Drinking Water: Estimating Uptake through the Skin and by Inhalation*. S. S. Olin. Washington, D.C., CRC Press/ International Life Sciences Institute Press.

Mills, Christina J., Richard J. Bull, Kenneth P. Cantor, John Reif, Steve E. Hrudey and Patricia Huston, 1998. Health Risks of Drinking Water Chlorination By-Products: Report of an Expert Working Group. *Chronic Diseases in Canada* **19**(3): 91-102.

Morris, Robert D., Anne-Marie Audet, Italo F. Angelillo, Thomas C. Chalmers and Frederick Mosteller, 1992. Chlorination, Chlorination By-Products, and Cancer: A Meta-Analysis. *American Journal of Public Health* **82**(7): 955-963.

Murray, F.J., B.A. Schwetz, J.G. McBride and R.E. Staples, 1979. Toxicity of Inhaled Chloroform in Pregnant Mice and Their Offspring. *Toxicology and Applied Pharmacology* **50**: 515-522.

Narotsky, Michael G., Rex A. Pegram and Robert J. Kavlock, 1997. Effect of Dosing Vehicle on the Developmental Toxicity of Bromodichloromethane and Carbon Tetrachloride in Rats. *Fundamental and Applied Toxicology* **40**: 30-36.

Nieuwenhuijsen, Mark J., Mireille B. Toledano and Paul Elliott, 2000. Uptake of Chlorination Disinfection By-Products; A Review and a Discussion of its Implications for Exposure Assessment in Epidemiological Studies. *Journal of Exposure Analysis and Environmental Epidemiology* **10**: 586-599.

Paustenbach, Dennis J., 2000. The Practice of Exposure Assessment: A State-of-the-Art Review. *Journal of Toxicology and Environmental Health, Part B* **3**: 179-291.

Petrucci, Ralph H., 1989. *General Chemistry*. New York, Macmillan Publishing Company.

Reif, John S. , Maureen C. Hatch, Michael Bracken, Lewis B. Holmes, Bernard A. Schwetz and Philip C. Singer, 1996. Reproductive and Developmental Effects of Disinfection By-Products in Drinking Water. *Environmental Health Perspectives* **104**(10): 1056-1061.

Richardson, Susan D., 1998. Drinking Water Disinfection By-Products. in *Encyclopedia of Environmental Analysis and Remediation*. R. A. Meyers. New York, John Wiley & Sons: 1398-1421.

Rizak, Samantha, Steve E. Hrudey, Kelly Kjartanson, Trevor Satchwill and Kenneth L. Froese, 2000. *Distribution System Monitoring Data as a Basis to Estimate Human Exposure to Halogenated Disinfection By-products*. Ninth

National Conference on Drinking Water, Ottawa, Ontario, Canadian Water and Wastewater Association.

Rook, J.J., 1974. Formation of Haloforms During Chlorination of Natural Waters. *Water Treatment and Examination* **23**: 234-243.

Rossman, Lewis A., Richard A. Brown, Philip C. Singer and John R. Nuckols, 1999. *Kinetics of THM and HAA Production in a Simulated Distribution System.* Water Quality Technology Conference, Tampa, Florida, American Water Works Association.

Rothman, Kenneth J. and Sander Greenland, 1998. *Modern Epidemiology.* Philadelphia, Lippincott-Raven Publishers.

Ruddick, J.A., D.C. Villeneuve and I. Chu, 1983. A Teratological Assessment of Four Trihalomethanes in the Rat. *Journal of Environmental Science and Health* **B18**(3): 333-349.

Sawyer, Clair N., Perry L. McCarty and Gene F. Parkin, 1994. *Chemistry for Environmental Engineering.* New York, McGraw-Hill, Inc.

Schwetz, B.A., B.K.J. Leong and P.J. Gehring, 1974. Embryo- and Fetotoxicity of Inhaled Chloroform in Rats. *Toxicology and Applied Pharmacology* **28**: 442-451.

Shimokura, Gayle H., David A. Savitz and Elaine Symanski, 1998. Assessment of Water Use for Estimating Exposure to Tap Water Contaminants. *Environmental Health Perspectives* **106**(2): 55-59.

Singer, Philip C., 1999. *Variability and Assessment of Disinfection By-Product Concentrations in Water Distribution Systems.* The Safety of Water Disinfection: Balancing Chemical and Microbial Risks, Miami Beach, Florida, International Life Sciences Institute.

Smith, M.K., J.L. Randall, E.J. Read and J.A. Stober, 1989a. Teratogenic Activity of Trichloroacetic Acid in the Rat. *Teratology* **40**: 445-451.

Smith, M.K., J.L. Randall, E.J. Read and J.A. Stober, 1992. Developmental Toxicity of Dichloroacetate in the Rat. *Teratology* **46**: 217-223.

Smith, M.K., J.L. Randall, J.A. Sotber and E.J. Read, 1989b. Developmental Toxicology of Dichloroacetonitrile: A By-Products of Drinking Water Disinfection. *Fundamental and Applied Toxicology* **12**: 765-772.

Smith, M.K., J.L. Randall, D.R. Tocco, R.G. York, J.A. Stober and E.J. Read, 1988. Teratogenic Effects of Trichloroacetonitrile in the Long-Evans Rat. *Teratology* **38**: 113-120.

Snoeyink, Vernon L. and David Jenkins, 1980. *Water Chemistry*. New York, John Wiley and Sons, Inc.

Stevens, Alan A., Leown A. Moore and Richard J. Miltner, 1989. Formation and Control of Non-Trihalomethane Disinfection By-Products. *Journal of the American Water Works Association* **81**(8): 54-60.

Swan, Shanna H. and Kirsten Waller, 1998. Disinfection By-Products and Adverse Pregnancy Outcomes: What Is the Agent and How Should it Be Measured? *Epidemiology* **9**(5): 479-481.

Thompson, D.J., S.D. Warner and V.B. Robinson, 1974. Teratology Studies on Orally Administered Chloroform in the Rat and Rabbit. *Toxicology and Applied Pharmacology* **29**: 348-357.

Toth, G.P., K.C. Kelty, E.L. George, E.J. Read and M.K. Smith, 1992. Adverse Male Reproductive Effects following Subchronic Exposure of Rats to Sodium Dichloroacetate. *Fundamental and Applied Toxicology* **19**: 57-63.

U.S. EPA, 1992a. *Dermal Exposure Assessment: Principles and Applications*. Washington, D.C. U.S. EPA, Office of Health and Environmental Assessment, Office of Research and Development. EPA/600/8-91/011.

U.S. EPA, 1992b. Guidelines for Exposure Assessment. *Federal Register* **57**(104): 22888-22938.

U.S. EPA, 1997. *Exposure Factors Handbook, Volume I, General Factors*. Washington, D.C. U.S. EPA. 1997. EPA/600/P-95/002Fa.

U.S. National Research Council, 1987. Biomarkers in Environmental Health Research. *Environmental Health Perspectives* **74**: 3-9.

U.S. National Research Council, 1989. *Biologic Markers in Reproductive Toxicology*. Washington, D.C., National Academy Press.

U.S. National Research Council, 1991. *Human Exposure Assessment for Airborne Pollutants: Advances and Opportunities*. Washington, D.C. National Research Council.

U.S. Public Health Service, 1964. *Smoking and Health: Report of the Advisory Committee to the Surgeon General of the Public Health Service*. Washington, D.C. U.S. Public Health Service. PHS Publication No. 1103.

Valentine, Richard L., Brian Anferman, Shelley Hackett, Peter Vikesland and Skipp Slattenow, 1999. *Characterization of Disinfectant Decay and DBP Formation in the Presence of Water Distribution System Deposits*. Water Quality Technology Conference, Tampa, Florida, American Water Works Association.

Wallace, Lance, Edo Pellizzari, Ty Hartwell, Harvey Zelon, Charles Sparacino, Rebecca Perritt et al., 1986. Concentrations of 20 Volatile Organic Compounds in the Air and Drinking Water of 350 Residents of New Jersey Compared with Concentrations in their Exhaled Breath. *Journal of Occupational Medicine* **28**(8): 603-608.

Wallace, L.A., 1997. Human Exposure and Body Burden for Chloroform and Other Trihalomethanes. *Critical Reviews in Environmental Science and Technology* **22**(2): 113-194.

Wallace, Lance A., Edo Pellizzari, Ty Hartwell, Martin Rosenzweig, Mitchell Erickson, Charles Sparacino et al., 1984. Personal Exposure to Volatile Organic Compounds I. Direct Measurements in Breathing-Zone Air, Drinking Water, Food, and Exhaled Breath. *Environmental Research* **35**: 293-319.

Weisel, Clifford and Wan-Kuen Jo, 1996. Ingestion, Inhalation, and Dermal Exposures to Chloroform and Trichloroethene from Tap Water. *Environmental Health Perspectives* **104**(1): 48-51.

Weisel, Clifford P., Hekap Kim, Patricia Haltmeier and Judith B. Klotz, 1999a. Exposure Estimates to Disinfection By-Products of Chlorinated Drinking Water. *Environmental Health Perspectives* **107**(2): 103-110.

Weisel, Clifford P., John C. Little, Nancy Chiu, Spyros N. Pandis, Cliff Davidson and Charles R. Wilkes, 1999b. Developing Exposure Estimates. in *Exposure to Contaminants in Drinking Water: Estimating Uptake through the Skin and by Inhalation*. S. S. Olin. Washington, D.C., CRC Press; International Life Sciences Institute.

Wilcosky, Timothy C., 1990. Criteria for Selecting and Evaluating Markers. in *Biological Markers in Epidemiology*. B. S. Huylka, T. C. Wilcosky and J. D. Griffith. New York, NY, Oxford University Press.

Wilkins, J.R. and G.W. Comstock, 1981. Source of drinking water at home and site-specific cancer incidence in Washington County, Maryland. *American Journal of Epidemiology* **114**: 178-190.

World Health Organization, 1993. *Biomarkers and Risk Assessment: Concepts and Principles*. Finland, World Health Organization.

Young, T.B., D.A. Wolf and M.S. Kanarek, 1981. Epidemiologic study of drinking water chlorination and Wisconsin female cancer mortality. *Journal of the National Cancer Institute* **67**: 1191-1198.

Zierler, Sally, Lisa Feingold, Robert A. Danley and Gunther Craun, 1988. Bladder Cancer in Massachusetts Related to Chlorinated and Chloraminated Drinking Water: A Case-Control Study. *Archives of Environmental Health* **43**(2): 194-200.

Chapter 2

Exposure Assessment Framework

To begin to answer the questions posed in the last section, we require an approach for comparing epidemiology studies, based on their exposure assessment. To this end, a framework was developed that divides exposure assessment into three basic aspects, outlined below.

1. Personal information vs. Aggregated population information

This aspect of exposure assessment pertains to information on water supply, water consumption, water use, and other exposures. It is one element in determining how accurate the exposure classification is at an individual level in an epidemiology study. The information described in this section is termed either personal or aggregated population information for the purposes of the exposure comparison model. Personal information is obtained from interviews or questionnaires including questions regarding the water supply[1], water consumption[2], water use[3], and other exposures[4]. Depending on the time lapse between the occurrence of the interview and the time of the exposure of interest, personal data can be subject to recall bias (Fenster et al., 1992). Personal interviews also provide useful information as to place of maternal residence at various times during pregnancy and migration of residence during pregnancy.

Aggregated population data generally provides information only on water supply. With aggregated population data, the subject's residence is linked to a geographic area served by a particular water utility or type of water. For example, a subject's maternal residence from birth records can be linked to a particular water utility providing chlorinated surface water and reporting certain levels of THMs. The exposure for this subject is then determined to be the THM concentrations reported by the providing water utility. No interviews are carried out or questionnaires filled in. This type of data assumes that the location of residence defines the water supply used. It can not take into account alternate water supplies such as private wells or use of bottled water. It cannot account for consumption or use of water either at the maternal residence or in locations different from the maternal residence. It cannot account for alternate routes of exposure to agents in chlorinated drinking water. Aggregated population data makes the assumption that there are no temporal and spatial differences in DBP concentrations between the distribution sampling point and date of sampling reported by the water utility and the subject's home during the exposure time of interest. In addition, it is assumed that the concentrations reported in monitoring records represent the actual exposure of the study subjects. Aggregated population data can also include the use of national statistics on water consumption and water use habits, although these data have not been used in epidemiologic studies to date.

---

[1] Do you obtain your water from a public water supply or a private well? Which water utility supplies your water?

[2] How many glasses of cold tap water do you drink per day? How many glasses of beverages made with cold or hot tap water do you drink per day? How much bottled water do you usually drink per day?

[3] How often and for how long do you shower and/or bathe per week? How often and for how long do you bathe others, such as infants or invalids? How often and for how long do you wash dishes and/or clothes by hand?

[4] How often and for how long do you swim in a chlorinated swimming pool per week? How often and for how long do you use a Jacuzzi per week?

2. *Specific vs. non-specific chemicals*

The next aspect of exposure assessment is the investigation of specific chemicals or agents *vs*. more general water characteristics. Data on specific agents includes the measurement of concentrations of individual compounds or groups of compounds. Two examples of specific data include monitoring data on concentrations of total THMs in a distribution system, and individual THM concentrations from sampling in subjects' homes. Examples of non-specific data or general information on water characteristics include type of water source or disinfection method. A study that investigates the differences in surface water vs. ground water or chlorine disinfection vs. no disinfection rather than measured DBP concentrations is using non-specific data. In order to determine a link between a causal agent (e.g. a DBP or DBPs) and an effect (adverse reproductive outcomes) it is necessary to measure individual exposure to the causal agent. Although the causal agent(s) for reproductive effects is (are) not known yet, the DBPs measured in past studies may yet prove to be causal agents or may be correlated to causal agents identified in the future. Therefore, measurement of specific DBPs in drinking water provides a more accurate assessment of individual exposure than measurement of non-specific water characteristics and may aid in determining the causal agent(s).

3. *Direct measurement vs. indirect measurement*

This aspect of exposure assessment refers to the second element of exposure assessment, the measurement of specific vs. non-specific compounds, and applies only to the measurement of specific compounds. Classifying the measurement of specific compounds as direct or indirect is an indication of the temporal and spatial distance between the point of measurement and the point of exposure for individual subjects. For example, information on concentrations of specific compounds may be obtained from water utilities. This information is usually temporally and spatially removed from the exposure point of the study subjects (their home or work place) and as such is an indirect measure of the concentration of DBPs for the purpose of individual exposure assessment. However, if samples are taken from a subject's home or workplace during the exposure time of interest and analysed for specific DBP concentration, those data are then a more direct measurement of individual exposure.

Having broken down exposure assessment into several key elements, we can develop an approach for classifying epidemiology studies looking at a possible association between drinking water and adverse reproductive outcomes in terms of their exposure assessment. The framework for comparison consists of five categories, based on various combinations of the key aspects of exposure assessment previously discussed. The categories are presented graphically in Figure 2-1 and present a trend from least accurate to most accurate individual exposure.

Figure 2-1. Framework for Categories of Exposure Assessment Used in Reproductive Epidemiological Studies

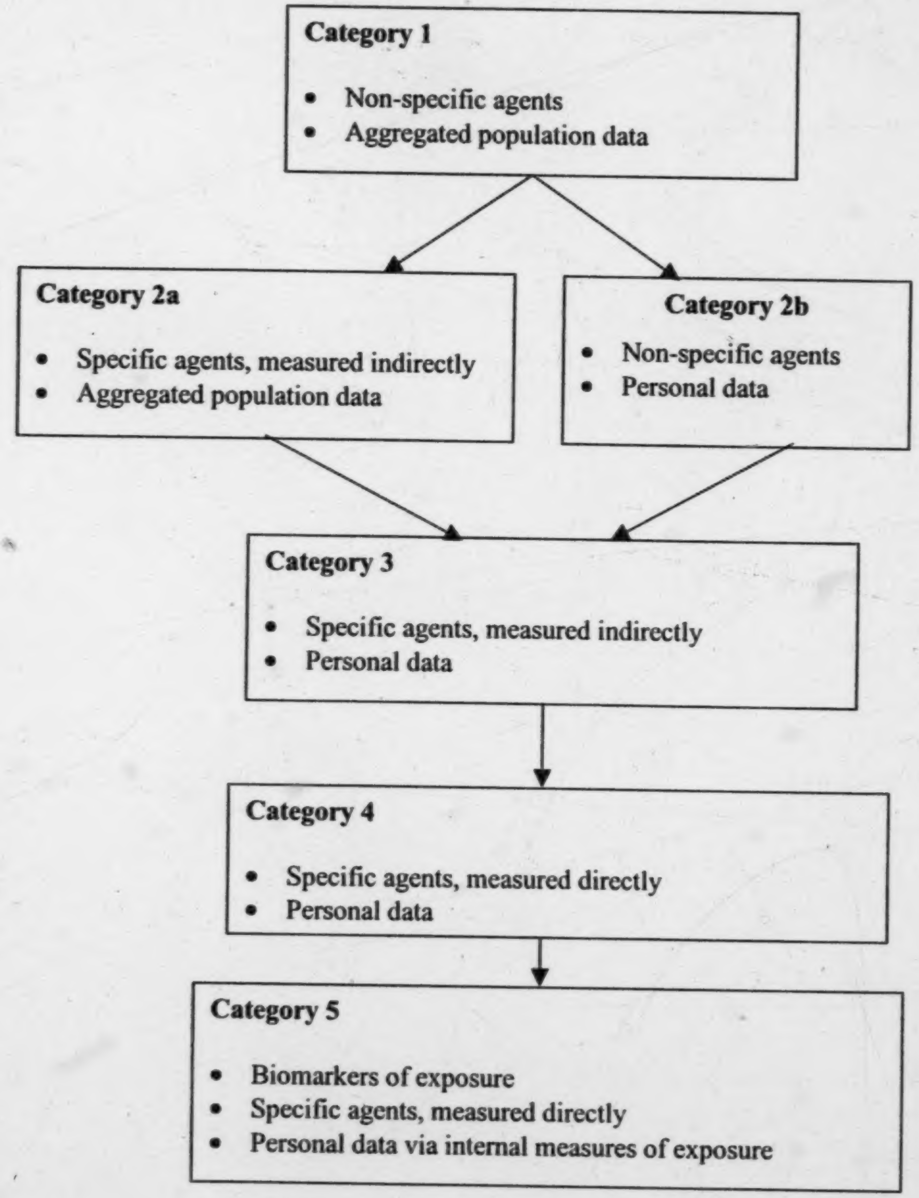

Category 1
Non-specific agents
Aggregated population data
Category 2a
Specific agents, measured indirectly
Aggregated population data
Category 2b
Non-specific agents
Personal data
Category 3
Specific agents, measured indirectly
Personal data
Category 4
Specific agents, measured directly
Personal data
Category 5
Biomarkers of exposure
Specific agents, measured directly
Personal data via internal measures of exposure

Table 2-1. Summary of epidemiology studies investigating a link between drinking water and various adverse reproductive effects, sorted by exposure category

Category	Study	Exposure	Outcome(s) investigated	Reported Significant Results (OR)
Category 1: Non-specific agents measured indirectly; ecological data	Tuthill et al., 1982 (retrospective ecological)	**Non-specific agents:** chlorine vs. chlorine dioxide disinfection **Aggregated population data:** Exposure classified by subjects' residence in community identified by type of disinfection	Premature delivery, weight loss after birth	No ORs calculated; statistically significant positive association with exposure to chlorine dioxide treated water during pregnancy
	Aschengrau et al., 1993 (case-control)	**Non-specific agents:** Water source (ground, surface, mixed) and disinfection type (chlorination or chloramination) **Aggregated population data:** Exposure assigned by maternal residence from medical records in first trimester or at pregnancy outcome in community identified by type of disinfection and water source	Respiratory defects Urinary tract defects Stillbirths	For chlorination *vs.* chloramination: OR 3.2 95% CI: 1.1-9.5 OR 4.1 95% CI 1.2-14.1 OR 2.6 95% CI 0.9-7.5
	Kanitz et al., 1996 (cross-sectional)	**Non-specific agents:** Disinfection with chlorine dioxide, sodium hypochlorite, both alternately, or no disinfection **Aggregated population data:** Exposure assigned by maternal residence from medical records at pregnancy outcome in community identified by type of disinfection	Smaller body length Smaller cranial circumference Neonatal jaundice	Chlorine dioxide vs. no disinfection (ND) OR 2.0 95% CI 1.2-3.3 Sodium hypochlorite vs. ND OR 2.3 95% CI 1.3-4.2 Chlorine dioxide vs. ND OR 2.2 95% CI 1.4-3.9 Sodium hypochlorite vs. ND OR 3.5 95% CI 2.1-8.5 Chlorine dioxide vs. ND OR 1.7 95% CI 1.1-3.1

Category	Study	Exposure	Outcome(s) investigated	Reported Significant Results (OR)
Category 1 continued	Magnus et al., 1999 (retrospective cohort)	**Non-specific agents:** Chlorination or no chlorination and measurement of colour **Aggregated population data:** Exposure assigned by maternal residence from medical records at pregnancy outcome in municipality identified by type of disinfection and amount of colour	All malformations Urinary tract defects	For high colour with chlorination *vs.* low colour and no chlorination: OR 1.14 95% CI 0.99-1.31 OR 1.99 95% CI 1.10-3.57
	Yang et al., 2000 (retrospective cohort)	**Non-specific agents:** Chlorination vs. no chlorination **Aggregated population data:** Exposure assigned by maternal residence from medical records at pregnancy outcome in municipality identified as chlorinating or non-chlorinating	Preterm delivery (<37 gestational weeks)	Chlorination *vs.* no chlorination OR 1.34 95% CI 1.15-1.56
	Kallen and Robert, 2000 (retrospective cohort)	**Non-specific agents:** Chlorine dioxide vs. sodium hypochlorite used as a disinfectant or no disinfection **Aggregated population data:** Exposure assigned by location of birth and maternal residence from public records in municipality identified as using chlorine dioxide, sodium hypochlorite, or no disinfection	Premature delivery (< 32 weeks gest.) Premature delivery (<37 weeks gest.) Low birth weight Short body length (< 43 cm length) Short body length (<47 cm length) Very small head circumf. (<31cm)	Sodium hypochlorite *vs.* chlorine dioxide OR 1.22 95% CI 1.00-1.48 OR 1.09 95% CI 1.01-1.17 OR 1.15 95% CI 1.05-1.26 OR 1.97 95% CI 1.30-2.97 OR 1.25 95% CI 1.10-1.43 OR 1.46 95% CI 1.07-1.98

Category	Study	Exposure	Outcome(s) investigated	Reported Significant Results (OR)
	Jaakola et al., 2001 (retrospective cohort)	**Non-specific agents:** Chlorination or no chlorination and measurement of colour **Aggregated population data:** Exposure assigned by maternal residence from medical records at pregnancy outcome in municipality identified by type of disinfection and amount of colour	Low birth weight (<2500g) Small for gest. age (birth weight ≤10percentile) Preterm delivery (gest. age <37 weeks)	For high colour with chlorination *vs.* low colour and no chlorination: OR 0.97 95% CI 0.89-1.06 OR 1.00 95% CI 0.91-1.10 OR 0.91 95% CI 0.84-0.99 (protective effect)
Category 2a: Specific agents, measured indirectly; ecological data	Kramer et al., 1992 (case-control)	**Specific agents:** Individual THM species concentrations **Indirectly measured:** Concentration data from a 1987 municipal water survey **Aggregated population data:** Exposure assigned by municipality of residence determined by maternal residence information taken from public records from Jan. 1, 1989 to June 30, 1990	Intrauterine growth retardation	TCM≥10μg/L vs. undetectable TCM levels OR 1.8 95% CI 1.1-2.9

Category	Study	Exposure	Outcome(s) investigated	Reported Significant Results (OR)
	Bove et al., 1995 (cross-sectional)	**Specific agents:** TTHM concentrations **Indirectly measured:** Calculated monthly estimates of TTHM concentrations in distribution system from quarterly monitoring **Aggregated population data:** Exposure assigned by municipality of residence determined by maternal residence information taken from public records	Low birth weight (<2500g)	TTHMs >100ppb vs. ≤20ppb OR 1.42 50% CI[1] 1.22-1.65
			Small for gestational age	TTHM >100ppb vs. ≤20ppb OR 1.50 50% CI 1.36-1.65
			All birth defects	TTHM >80ppb vs. ≤20ppb OR 1.57 50% CI 1.42-1.75
			Central nervous system defects	TTHM >80ppb vs. ≤20ppb OR 2.59 50% CI 2.05-3.28
			Neural tube defects	TTHM >80ppb vs. ≤20ppb OR 2.96 50% CI 2.00-4.39
			Oral cleft defects	TTHM>100pp b vs. ≤20ppb OR 3.17 50%CI 2.05-4.89
			Cardiac defects	TTHM>80ppb vs. ≤20ppb OR 1.83 50% CI 1.38-2.43
			All cardiac defects	TTHM>80ppb vs. ≤20ppb OR 1.44 50% CI 1.23-1.68

---

[1] The 50% CIs are taken directly from the published study and are in contrast to all other studies that quote 95% CIs. The use of 50%CIs calls into question the reliability of these data.

Category	Study	Exposure	Outcome(s) investigated	Reported Significant Results (OR)
Category 2a: Specific agents, measured indirectly; ecological data	Gallagher et al., 1998 (retrospective cohort)	**Specific agents:** TTHM concentrations **Indirectly measured:** Data obtained from records of quarterly sampling in distribution systems and hydraulically modeled to determine "exposure levels" for census block groups **Aggregated population data:** Exposure assigned by residence in census block groups determined by maternal residence information taken from public records	Term low birth weight (≥37 weeks gestation and ≤5 lbs., 8 oz.) Low birth weight (≤5 lbs., 8 oz.)	TTHM≥61 μg/L vs. ≤ 20 μg/L OR 5.9 95% CI 2.0-17.0 TTHM≥61 μg/L vs. ≤ 20 μg/L OR 2.1 95% CI 1.0-4.8
	Klotz and Pyrch, 1999 (case-control)	**Specific agents:** Individual THM and TTHM concentrations **Indirectly measured:** Data obtained from records of quarterly sampling in distribution systems for time of exposure and one year later **Aggregated population data:** Exposure assigned by residence in geographic area served by a particular water utility determined by maternal residence information taken from public records	Neural tube defects	TTHMs 40+ppb (highest tertile *vs.* lowest tertile) OR 2.1 95% CI 1.1-4.0 (analysis restricted to isolated cases and known residence at conception) OR 1.7 95% CI 0.9-3.0 (restricted to isolated defects) OR 1.6 95% CI 0.9-2.70 (unrestricted analysis)
	Dodds et al., 1999 (retrospective cohort)	**Specific agents:** TTHM concentrations **Indirectly measured:** Data obtained from records of quarterly sampling in distribution systems **Aggregated population data:** Exposure assigned by residence in geographic area served by a particular water utility determined by maternal residence information taken from public records	Stillbirths	TTHM≥100μg/L vs. ≤49 μg/L OR 1.69 95% CI 1.1-2.59 (crude) TTHM≥100μg/L vs. ≤49 μg/L OR 1.66 95% CI 1.09-2.52 (adjusted)

Category	Study	Exposure	Outcome(s) investigated	Reported Significant Results (OR)
Category 2a continued	King et al., 2000 (retrospective cohort)	**Specific agents:** TTHM and individual THM concentrations **Indirectly measured:** Data obtained from records of quarterly sampling in distribution systems **Aggregated population data:** Exposure assigned by residence in geographic area served by a particular water utility determined by maternal residence information taken from public records	Stillbirths	TTHM ≥100 μg/L vs. <50 μg/L OR 1.66 95% CI 1.09-2.54 (adjusted) TCM ≥100 μg/L vs. <50 μg/L OR 1.56 95% CI 1.04-2.34 (adjusted) BDCM ≥20 μg/L vs. <5 μg/L OR 1.98 95% CI 1.23-3.49 (adjusted) BDCM per 10 μg/L OR 1.29 95% CI 1.10-1.53 (adj.)
			Asphyxia-related stillbirths	TTHM ≥100 μg/L vs. <50 μg/L OR 4.57 95% CI 1.93-10.77 (adjusted) TTHM 50-74 μg/L vs. <50 μg/L OR 2.67 95% CI 1.16-6.14 (adjusted) TCM ≥100 μg/L vs. <50 μg/L OR 3.15 95% CI 1.64-6.03 (adjusted) BDCM 10-19 μg/L vs. <5 μg/L OR 2.32 95% CI 1.18-4.55 (adjusted)

Category	Study	Exposure	Outcome(s) investigated	Reported Significant Results (OR)
	Dodds and King, 2001 (retrospective cohort)	**Specific agents:** Individual THM concentrations (TCM, BDCM) **Indirectly measured:** Data obtained from records of quarterly sampling in distribution systems **Aggregated population data:** Exposure assigned by residence in geographic area served by a particular water utility determined by maternal residence information taken from public records	Neural tube defects Chromosomal abnormalities Cardiovascular anomalies	BDCM ≥20 μg/L vs. <5 μg/L OR 2.5 95% CI 1.2-5.1 (adj.) TCM 75-99 μg/L vs. < 50 μg/L OR 1.9 95% CI 1.1-3.3 (adj.) BDCM ≥20 μg/L vs. <5 μg/L OR 0.3 95% CI 0.2-0.7 (adj.)* *protective effect
Category 2b: Non-specific agents measured indirectly; personal data	Hertz-Picciotto et al., 1989 (nested case-control)	**Non-specific agents:** Tap water *vs.* bottled water **Personal data:** Interview questions on bottled and tap water consumption in first 3 months of pregnancy at home and at work; qualitative question: tap water consumed more or less than bottled water; no quantitation of water consumption	Spontaneous abortion	Only or mostly tap water *vs.* only or mostly bottled water OR 1.7 95% CI 1.2-2.3 (crude) Any tap water *vs.* only bottled water OR 2.0 95% CI 1.1-3.7 (crude)
	Deane et al., 1992 (retrospective cohort)	**Non-specific agents:** Chlorinated tap water *vs.* bottled water **Personal data:** Interview data on amount of cold tap water consumed during pregnancy and for 3 months prior to pregnancy; any bottled water consumed at home; same questions for consumption at work	Spontaneous abortion	OR 3.4 95% CI 0.6-19.4 for any tap water consumption controlled for bottled water OR 2.2 95% CI 1.2-4.0 per category change when consumption of tap water treated as a categorical variable with three categories Protective effect of bottled water consumption controlled for tap water: OR 0.6 95% CI 0.1-2.6

Category	Study	Exposure	Outcome(s) investigated	Reported Significant Results (OR)
	Windham et al., 1992 (case-control)	**Non-specific agents:** Chlorinated tap water *vs.* bottled water **Personal data:** Interview data from cases on entire pregnancy and from controls on first 20 weeks of pregnancy; number of glasses of cold tap water or beverages made from cold tap water per day at work and at home; question re. change in consumption during pregnancy compared to before pregnancy; subsection of cases and control asked re. showering and bathing time per week	Spontaneous abortion	1-2 glasses/day vs. 0 glasses/day OR 1.3 95% CI 1.0-1.7 (crude) OR 1.3 95% CI 0.95-1.7 (adjusted) ≥3 glasses/day vs. 0 glasses/day OR 1.2 95% CI 0.74-1.5 (crude) OR 1.2 95% CI 0.92-1.5 (adjusted) Any (≥ 0.5 glasses/day) vs. 0 glasses/day OR 1.2 95% CI 1.0-1.5 (crude) OR 1.2 95% CI 0.98-1.6 (adj.) Protective effect of bottled water: OR 0.79 95%CI 0.65-0.96
Category 2b continued	Fenster et al., 1992 (case-control)	**Non-specific agents:** Chlorinated tap water *vs.* bottled water **Personal data:** First interview: same as Windham et al., 1992; second interview: abbreviated form of first interview	Spontaneous abortion	Any tap water *vs.* no tap water: OR 0.71 95% CI 0.43-1.2 (first interview) OR 1.1 95% CI 0.69-1.9 (second interview)

Category	Study	Exposure	Outcome(s) investigated	Reported Significant Results (OR)
	Wrensch et al., 1992 (case-control)	**Non-specific agents:** Chlorinated tap water *vs.* bottled water **Personal data:** Interview questions: how many glasses of cold tap water consumed at home during pregnancy or in 3 months before pregnancy; any change in amount during first 3 months of pregnancy, if so how much; same questions about tap water consumption at work; any bottled water consumed at home during or in 3 months before pregnancy; did usually drink bottled water at home; same questions about bottled water consumption at work	Spontaneous abortion	Any tap water *vs.* no tap water: OR 4.0 95% CI 1.8-9.1 (crude) OR 6.9 95% CI 2.7-17.7 (adjusted) Protective effect: any bottled water vs. no bottled water OR 0.26 95% CI 0.16-0.43
	Swan et al., 1998 (prospective cohort)	**Non-specific agents:** Chlorinated tap water *vs.* bottled water **Personal data:** Interview questions: how many glasses of water consumed in week beginning with last menstrual period and week before the interview; if consumption differed, time change occurred determined; any treatment of tap water before consumption(water filter use, let it stand, etc.); number of showers per week and average length	Spontaneous abortion	≥6 glasses cold tapwater/day *vs.* none: OR 2.17 95% CI 1.22-3.87 High tap water and no bottled water *vs.* low tap water and high bottled water: OR 4.58 95% CI 1.97-10.64 High bottled water and no tap water vs. some tapwater and no bottled water (protective effect): OR 0.22 95% CI 0.09-0.51

Category 3: Specific agents measured indirectly; personal data	Savitz et al., 1995 (case-control)	**Specific agent:** TTHM concentrations **Indirectly measured:** Concentration data obtained from records of quarterly sampling in distribution systems **Personal data:** Interview questions on primary source of drinking water in home; # of glasses of tap water consumed per day around the time of pregnancy	Spontaneous abortion	For highest sextile of dose (TTHM conc. x amount of water consumed vs. lowest sextile dose) OR 1.7 95% CI 1.1-2.7 *In general, an increase in the amount of water consumed resulted in a decreased risk for the outcomes of interest.
	Waller et al., 1998 (prospective cohort)	**Specific agents:** Concentrations of individual THMs and TTHMs **Indirectly measured:** Concentration data obtained from records of quarterly sampling in distribution systems; averaged all distribution system measurements taken by water utility in the first trimester of pregnancy for subjects **Personal data:** Interview questions: number of glasses of water consumed in week beginning with last menstrual period and week before the interview; if consumption differed, time change occurred determined; treatment of tap water before consumption (water filter use, let it stand, etc.); number and average length of showers per week	Spontaneous abortion	≥5 glasses cold tap water/day with ≥75 μg/L TTHM vs. <5 glasses/day with <75 μg/L TTHM OR 1.8 95% CI 1.1-3.0 ≥5 glasses cold tap water/day with ≥18 μg/L BDCM vs. <5 glasses/day with <18 μg/L BDCM OR 2.0 95% CI 1.2-3.5 (BDCM alone) OR 3.0 95% CI 1.4-6.6 (adjusted for other THMs)
Category 4: Specific agents measured directly; personal data	Klotz and Pyrch, 1999 (case-control)	**Specific agents:** Concentrations of individual THMs, TTHMs, HANs, HAAs **Directly measured:** Concentration data sampled in subjects' home taps one year after exposure time of interest **Personal data:** Interview data: quantities of hot and cold beverages made using tap water consumed per day, length and frequency of bathing/showering, exposure in swimming pools; in the 3 months before and the first trimester of pregnancy	Neural tube defects	TTHMs ≥40ppb vs. <5ppb OR 1.9 95% CI 1.0-4.0 (restricted to isolated defects) OR 1.7 95% CI 0.9-3.8 (unrestricted analysis)

Category 1 studies provide the least accurate individual exposure assessment. The key characteristics of Category 1 studies are the use of non-specific agents measured indirectly, and aggregated population-based information. Category 1 studies generally investigate the association between a particular water source, such as surface vs. ground water, or a type of disinfection, for example chlorination vs. chloramination, and the outcomes of interest. The information linking subjects to the exposure is aggregated rather than individual. Subjects' exposure classification is determined by their location of residence in relation to the water characteristic of interest in a geographic area.

Category 2 studies take the exposure assessment one step further. There are two sub-classes of Category 2 studies: 2a and 2b. Both are improvements on Category 1 studies in terms of exposure assessment in that they advance one of the key aspects of exposure assessment.

Category 2a studies improve on Category 1 exposure assessment with the use of data for specific agents. For example, instead of investigating types of disinfection as in a Category 1 study, a Category 2a study looks at concentrations of groups of DBPs or specific DBPs. By looking at differing levels of DBP concentrations, studies in this category can establish a crude dose-response relationship. However, the specific DBP data in a Category 2a study is taken from water utility records and is therefore indirectly measured. The link to the study subjects is aggregated, based on a geographical link between the subject's residence obtained from public records (usually hospital birth records) and the measured characteristics of the water obtained from the providing water utility.

Category 2b studies are similar to Category 1 studies in that they investigate non-specific agents, such as water source or disinfection method. However, the improvement from Category 1 studies comes from the application of personal exposure data. In other words, the subjects are contacted directly regarding their water source and consumption habits. Studies in this category can quantify the amount of water consumed by individual subjects and can therefore establish a crude dose-response relation in terms of consumption. In addition, information on alternate exposures from water use such as showering and bathing can be obtained.

The next category of studies combines the improvements of Categories 2a and 2b to determine individual exposure even more accurately. Category 3 studies use personal data from interviews and questionnaires to determine the water supply and water use habits of individual subjects. In addition, specific agents such as concentrations of groups of compounds or individual compounds are quantified. However, the specific exposure data are once more indirectly measured since the data for the specific agents are obtained from water utility records. Category 3 studies are able to determine a more accurate dose-response relationship than Category 1 or Category 2 studies.

Category 4 studies have the same general elements as Category 3 studies in that they measure specific agents and involve the collection of personal data in the form of interviews or questionnaires. However, the major improvement in Category 4 studies over Category 3 studies is the direct measurement of the specific agents. The agents are measured at the point of exposure; in other words, water samples are taken from taps in the subject's home or workplace during the exposure time period of interest and the specific agents of interest are quantified. This category contains the elements of the most accurate exposure assessment possible, barring the use of biomarkers of exposure. Studies in this category can include information on alternate routes of exposure, including showering and bathing, used in conjunction with the direct measurement of the water concentrations of specific agents of interest. Only one study to date has incorporated elements of Category 4.

Category 5 studies utilize biomarkers of exposure to measure individual exposure. Biomarkers of exposure can be described in terms of the key aspects of exposure upon which this framework is based. Specific agents are measured. These agents may be DBPs of interest as in previous categories, or they may be the metabolites of the DBPs. The specific agents are measured directly as concentrations in the subject's biological tissues or fluids and as such are an internal measure of exposure. The exposure information is certainly personal as the biomarkers are measured in each individual. There are no Category 5 studies to date since suitable biomarkers of exposure have yet to be sufficiently validated.

The problems associated with biomarkers of exposure for DBPs were discussed in Chapter 1. In addition to the short half-lives of most DBPs in the human body, there are the issues of the time and effort required for sampling and analyzing the biomarkers, as well as the burden placed on the study subjects. Nevertheless, biomarkers are still considered the "gold standard" for assessing exposure to DBPs in treated drinking water (Swan and Waller, 1998).

With this framework in mind, it is possible to begin to compare the studies completed to date that have investigated an association between exposure to chlorinated drinking water and adverse reproductive effects. To start with, we consider the first of the questions asked with respect to these studies: How accurately have reproductive epidemiology studies to date characterized individual exposure to possible causal agents in drinking water? To begin to answer this question, it is necessary to describe the exposure assessment used in each of the studies, arranged in the categories described in the framework developed above.

Summary of Epidemiology Studies

The epidemiology studies to date that have investigated a possible link between agents in treated drinking water and adverse reproductive effects are described in detail below and are summarized in Table 2-1.

Category 1 Studies

These studies incorporated exposure assessments at the far end of our classification framework from the "gold standard" of biomarkers. The exposures were based on non-specific agents measured indirectly and aggregated population exposure information.

In a retrospective aggregated study using historical records from the 1940s, the hypothesis was tested that newborns with prenatal exposure to water treated by chlorine dioxide as a result of consumption by their mothers would have a higher risk of adverse reproductive outcomes[5] (Tuthill et al., 1982). Maternal residences from birth and hospital records were used to identify subjects in geographic areas served by water utilities using either chlorine dioxide or chlorine as disinfectants.

To determine why this study is classified as a Category 1 study, it is necessary to look at the key aspects of exposure assessment. The proposed causal agent is non-specific in that it is the type of disinfection being investigated and not concentrations of DBPs. The amount of chlorine or chlorine dioxide added to the water was not measured. However, a surrogate, residual chlorine levels at the water treatment plants, was used as an indicator of chlorine dioxide levels. This information was used to determine the time period of the study; the decision was made to investigate infants born in the one-year period following the year of the highest chlorine dioxide dosage in order to maximize the probability of seeing outcomes. Aggregated population data were used to determine the exposure status of study subjects. Maternal residences (obtained from birth records) were linked to the geographical areas served by the water utilities using one of the two types of disinfection.

A 1993 case-control study was carried out to investigate a link between a variety of adverse reproductive outcomes[6] and exposure to chlorinated vs. chloraminated water, in addition to other water quality parameters (Aschengrau et al., 1993). Subjects in this study were chosen from a previous study looking at behavioral factors and risks for late adverse pregnancy outcomes. They were grouped according to outcome into three case groups and one control group. The proposed causal agent investigated was the type of disinfection, a non-specific agent. Information was also obtained on the source of drinking water prior to treatment (surface water, ground water, or a mixed source). Aggregated population data was used to assign subjects to type of disinfection and water source by linking maternal residence either in the first trimester or at pregnancy outcome (from medical records) to drinking water data during that time frame from the utility(ies) in the town of residence.

---

[5] The outcomes of interest included premature delivery, mean birth weight, mean maximum weight loss, mean weight loss at 6 days of age, jaundice, and infant mortality.

[6] Congenital anomalies, still births and neonatal deaths

A cross-sectional study was done in Italy to investigate a link between methods of disinfecting drinking water and various general outcomes at birth[7] (Kanitz et al., 1996). Disinfection methods considered were chlorine dioxide, sodium hypochlorite, both used alternately, or no disinfection. Aggregated population data was used to identify exposed and non-exposed subjects. Subjects were identified as exposed to one of the types of disinfection or not exposed to any disinfection based on their geographic area of residence as determined from hospital records.

A retrospective cohort study in Norway was carried out to investigate a variety of birth defects[8] with respect to chlorination of drinking water containing organic compounds (Magnus et al., 1999). The study population included all children born in Norway between 1993 and 1995 to mothers living in municipalities in which the chlorination status of at least one water utility was known and for which an indication of the organic content of the water could be calculated. The agents of interest investigated were a combination of chlorination status and organic content of the water, as measured by colour, which was shown in Norwegian waters to correlate highly with the concentration of dissolved organic carbon. These parameters were chosen because they represent the two main contributing reagents in the formation reaction of disinfection by-products. However, these agents are still non-specific as actual concentrations of DBPs were not measured[9]. Individuals were assigned exposure status with aggregated population data linking information on maternal residence from the Norwegian Birth Registry to the municipality of residence. The calculation of an exposure score for any particular individual was complicated by the fact that many municipalities are served by more than one water utility, which may or may not chlorinate their water. A "chlorination proportion" was developed for all municipalities to give a proportion of the population in each municipality that was served by chlorinated water based on the number of utilities that chlorinate their water and the fraction of the population those utilities serve in each municipality. In addition, a "weighted colour mean" was developed for all municipalities based on the water colour (a surrogate for DOC) reported by each water utility and the fraction of the population served by that water utility. If colour was not reported by all water utilities in a municipality, the weighted colour mean for the whole municipality was based on information from water utilities that did report colour. Chlorination and colour data were taken from 1994 water utility records.

---

[7] Outcomes included birthweight, body length at birth, cranial circumference at birth and neonatal jaundice.

[8] The definition of birth defects in this study was based on the International Classification of Diseases codes, eighth revision

[9] This study and the Jaakola et al. (2001) study could almost be classified in Category 2a. However, for the purposes of this paper the strict criterion of quantitation of specific DBPs will be adhered to. As a result, these studies will remain in Category 1, with the comment that the exposure assessment is more detailed than that normally found in a Category 1 study.

A retrospective cohort study in Taiwan was done to investigate an association between chlorination of drinking water and various birth outcomes[10] (Yang et al., 2000). Subjects were grouped according to exposure status based on their location of residence (aggregated population data taken from a national birth registry) in a municipality identified as chlorinating or not chlorinating the drinking water. If a municipality was defined as chlorinating the water supply, then greater than 90% of the population was estimated to be served by chlorinated water. If a municipality was defined as not chlorinating the water supply, then less than 5% of the population was estimated to be served by chlorinated water. Data on chlorination status was obtained from the Taiwan Water Supply Corporation.

A retrospective cohort study from Sweden (Kallen and Robert, 2000) was undertaken to investigate an association between chlorination of drinking water and birth outcomes[11]. Three exposure states were investigated based on method of chlorination of the drinking water and included chlorination by chlorine dioxide, chlorination by sodium hypochlorite, and no chlorination. Exposure status was determined by residence in a municipality served by chlorinated or non-chlorinated drinking water, based on maternal residence obtained from aggregated population data from the Swedish Medical Birth Registry. Data on drinking water chlorination status and type of chlorination used was obtained from published reports on municipality drinking water treatment for each municipality. However, these data were available for only three years: 1985, 1989, and 1994. Births were selected for the study if they occurred in a municipality where the disinfection type was the same before and after the time of the birth. For example, if a birth occurred in a municipality in 1987 and the method of chlorination reported for that municipality in 1985 and 1989 was the same, then that birth was accepted for the study. Municipalities with only one water source or more than one water source with the same method of disinfection were selected for the study. Municipalities with several water sources using different methods of disinfection were excluded. Disinfection with sodium hypochlorite was only reported in the 1989 and 1994 reports and therefore was only included in this study for this time period.

The most recent Category 1 study to date is a follow-up of the Magnus et al. (1999) study. The same outcome data base[12] and method of determining exposure by looking at chlorination and water colour in municipalities and linking them to aggregated population data on maternal residence were used in this study as in the original study (Jaakola et al., 2001)[13]. The method of linking chlorination with organic matter in the water is meant to produce a surrogate measure of potential DBP concentrations in the water[14].

---

[10] Birth outcomes investigated were term low birth weight (≥37 weeks gestation and <2500g birth weight) and preterm delivery (<37 gestational weeks).

[11] Birth outcomes investigated were multiple births, gestational duration, birth weight in singletons, intrauterine growth retardation in singletons, head circumference at birth in singletons, body mass index at birth in singletons, infant survival up to the age of one, Apgar score, neonatal jaundice, congenital malformations, childhood cancer, and neonatal diagnosis of hypothyroidism.

[12] The outcomes investigated were birth weight, low birth weight, small for gestational age, and preterm delivery.

Category 2a studies

Studies in Category 2a use indirectly measured specific agents (concentrations of DBPs obtained from water utilities) and aggregated population data to investigate associations between drinking water and adverse reproductive outcomes.

A population-based, case-control study was conducted in Iowa to investigate exposure to THMs, specifically chloroform (TCM), in relation to various adverse birth outcomes[15] (Kramer et al., 1992). Individual THM concentrations, including TCM, chlorodibromomethane (CDBM), bromodichloromethane (BDCM), and bromoform (TBM) were obtained from a 1987 municipal water survey. The THM concentrations were determined for each municipality. Mention is not made in this study of how many samples were taken in each municipality and at what frequency samples were taken during the year of the survey. THM concentration categories were designated based on the concentrations found in each municipality. Categories for TCM and BDCM were defined as undetectable (assumed to be <1 μg/L), low (1-9 μg/L), and high (≥10 μg/L). Categories for CDBM and TBM were defined as undetectable (assumed to be < 1 μg/L), low (1-3 μg/L), and high (≥4 μg/L). Subjects were assigned exposure to one of these categories based on their municipality of residence, determined from maternal residence information taken from birth certificates. This aggregated population data came from public records from the time period January 1, 1989 to June 30, 1990.

A cross-sectional study was conducted in New Jersey investigating several birth outcomes[16] in relation to various organic contaminants in drinking water, including THMs (Bove et al., 1995). Outcome information was obtained from birth certificates or fetal death certificates for the period from January 1985 to December 1988. Concentrations of total THM (TTHM) were determined from monitoring data obtained from the Bureau of Safe Drinking Water for 49 water utilities in 75 towns included in the study for the years 1984 to 1988. Monitoring data consisted of tap water samples from at least four locations in the distribution system on a quarterly basis, analyzed for TTHMs. Individual THM concentrations were not available. Monthly estimates of TTHM concentrations were determined and assigned to each gestational month for each live birth and fetal death. The method of estimating monthly TTHM concentrations was not described in this study. Estimated monthly concentrations were averaged over the time periods of interest for the various outcomes. For example, concentrations were averaged over the first trimester to investigate birth defect outcomes and fetal death. Concentrations were not measured directly in subjects' homes. Subjects were assigned exposure status by linking the maternal residence obtained from public records (aggregated population data) to the estimated TTHM concentrations for each municipality. Numerous assumptions were made in estimating each subject's monthly exposures.[17] It should be noted that significant associations were found in this study between exposure to TTHMs and several of the outcomes investigated. However, the confidence intervals reported were for the 50% range, which calls into question the meaning of these data.

---

[13] See footnote 7.

[14] Municipalities with high chlorine and DOC content would be predicted to have high levels of DBPs while those with low chlorine and DOC will most likely have low levels of DBPs.

[15] Birth outcomes investigated were low birthweight, prematurity, and intrauterine growth retardation.

[16] The outcomes investigated included low birth weight among term births (≥37 weeks gestation and <2500g birth weight), small for gestational age, preterm births (< 37 gestational weeks), very low birth weight (<1500g), and birth weight among term births (≥37 and < 42 gestational weeks). Fetal deaths were included except for those due to therapeutic abortions or chromosomal abnormalities. Other outcomes of interest for both live births and fetal deaths included central nervous system defects including neural tube defects, oral cleft defects, major cardiac defects, ventricular septal defects, all cardiac defects, and all defects typically reported by surveillance systems.

A retrospective cohort study in Colorado investigated a possible association between exposure to TTHMs in the third trimester of pregnancy and various birth outcomes[18] (Gallagher et al., 1998). Specific individual THM concentrations were obtained from water district monitoring records and then summed to give TTHM concentrations. The monitoring data included quarterly tap water samples from four different locations in the distribution system of each water district. The sampling points, dates and TTHM concentrations were then coded into a geographic information system (GIS). Information on census block group boundaries and maternal residence at time of birth was also entered into the GIS. Standard hydraulic models were then used to determine THM "exposure levels" in the census block groups contiguous to each quarterly sampling point. The model was then used to assign quarterly TTHM concentrations to each census block group. This information was then correlated with birth records to assign an exposure "score" for each subject based on the estimated TTHM concentrations during the last trimester of pregnancy. Ultimately, exposure was based on an indirect (although quite sophisticated) measurement of TTHM concentrations linked to aggregated population data on maternal residence obtained from public records.

---

[17] 1) The TTHM concentrations analysed at any particular sample date and location represent the TTHM concentrations in the distribution system of the entire municipality for one month before and one month after the sampling date. 2) The mother's municipality of residence at birth or other birth outcome (e.g. the address listed on the birth or fetal death certificate) did not change during the course of the pregnancy. 3) The drinking water supplied to the mother's residence was from the municipality's public system (e.g. the mother's source of drinking water was not a private well). 4) During pregnancy, the mother consumed and used water exclusively from her home or other location within the same municipality as her home 5) Showering and bathing provided considerable exposure to TTHMs via inhalation and dermal routes

[18] Low birthweight, low term birthweight, and preterm delivery

A population-based case-control study was carried out in New Jersey (Klotz and Pyrch, 1999) to investigate a possible association between exposure to DBPs and neural tube defects. This study presents an interesting case in this classification scheme since part of the study involved Category 2a exposure assessment, while another part of the study involved Category 4 exposure assessment. For that reason, this study will be discussed in both categories. Cases were identified through the New Jersey Birth Defects Registry records from 1993 and 1994. Approximately 10 potential controls were selected randomly for each month of 1993 and 1994 without regard for geographic location or case characteristics. Information on THM concentrations was obtained from water utility monitoring reports. THM monitoring data used included quarterly samples from the distribution systems of the water utilities analyzed for individual and total THMs. Study subjects were assigned THM exposure values based on their location of residence from ecological data taken from public monitoring records in relation to the geographic areas served by the various water utilities. The use of specific but indirectly measured and aggregated population data classifies this part of this study as Category 2a.

Three retrospective cohort studies in Nova Scotia (Dodds et al., 1999; King et al., 2000; Dodds and King, 2001) round out Category 2a. These studies used the same exposure and subject identification data to look at different adverse reproductive outcomes. The Dodds et al. (1999) study investigated a possible association between TTHM levels and adverse birth outcomes[19]. Subjects were identified using the Nova Scotia Atlee Perinatal Database for the years 1988 to 1995, inclusive. TTHM concentrations were obtained from the Nova Scotia Department of the Environment for the years 1987 to 1995, inclusive. TTHM concentrations are monitored by each public water facility in Nova Scotia by taking samples from at least three locations in the distribution system of each facility four times a year at irregular time intervals. TTHM levels were not directly measured in the subjects' homes or places of work. Linear regression was used on these monitoring data to estimate monthly TTHM concentrations for each facility. The aggregated population data on maternal residence obtained from the birth registry was linked to the geographic area served by each water facility. Using the monthly TTHM levels calculated by linear regression and information on gestation from the birth registry, TTHM exposure levels for various stages of pregnancy were estimated for all subjects. TTHM exposure levels for the last 3 months of pregnancy were estimated for outcomes related to fetal growth and time of delivery; for the first 2 months of pregnancy for outcomes such as cleft defects and cardiac defects; for neural tube defects the TTHM levels for one month before and one month after conception were used; for chromosomal abnormalities the TTHM levels in the 3 months before pregnancy were used; and TTHM values during the entire pregnancy were used to analyze for stillbirths.

---

[19] low birth weight (<2500 g), very low birth weight (<1500g), preterm delivery (<37 weeks gestation), small for gestational age, and stillbirths. Other outcomes investigated included congenital anomalies such as neural tube defects, cleft lip and palate, major cardiac defects, and chromosomal abnormalities.

Exposures were categorized based on average TTHM concentrations over the specified time periods and were as follows: 0-49 μg/L (referent category), 50-74 μg/L, 75-99 μg/L, ≥100 μg/L.

The second retrospective cohort study in Nova Scotia (King et al., 2000) used the same databases for subjects and THM concentrations as the first (Dodds et al., 1999). However, the second study looked not only at TTHM levels but individual THM levels as well. In addition, the only outcome analyzed was stillbirths. Concentrations of individual THMs were subjected to a least-squares regression to estimate monthly concentrations for each water facility. These concentrations were assumed to be the concentrations in the geographic areas served by each water facility. The monthly THM concentrations in each geographic area were then linked to the aggregated population data on maternal residence from the birth registry and the THM levels were averaged over the entire pregnancy. It was assumed that the mother's residence at birth was her residence during the entire pregnancy. The exposure data were analysed both as continuous data looking at per 10 μg/L changes and as categorical data using the same categories as the Dodds et al. (1999) study.

The third Nova Scotia study (Dodds and King, 2001) used the same databases, methods, and assumptions as the second of the Nova Scotia studies (King et al., 2000) to identify subjects and exposures. In this study, exposure to TCM and BDCM was investigated in relation to birth defects[20]. Other THM compounds were not of high enough concentrations to be included in this analysis. THM concentrations from public water utility monitoring data were subjected to a least-squares regression to estimate monthly concentrations for each water facility and the geographical area served by each water facility was linked to the aggregated population data on maternal residence. The monthly concentration data for different time periods were determined, specific to each outcome[21]. The concentration categories for TCM were < 50 (referent), 50-74, 75-99, and ≥ 100 μg/L. The concentration categories for BDCM were < 5 (referent), 5-9, 10-19, and ≥20 μg/L.

Category 2b studies

The studies in Category 2b improve exposure assessment over Category 1 studies in that subjects are contacted individually about their water consumption and water use habits. However, non-specific agents are investigated and actual concentrations of DBPs are not measured. Many of the studies in this category directly follow up the study mentioned earlier from Santa Clara County, California in 1980-1981 (Deane et al., 1989). Ironically, this study did not look specifically at DBPs in drinking water, but at a water supply that had been contaminated with industrial solvents. The results of this study suggested that alternate factors inherent in chlorinated drinking water could have a causal association with spontaneous abortions. This study launched a number of studies looking at a causal association between treated drinking water and adverse reproductive effects using a whole range of exposure assessment techniques.

---

[20] Including neural tube defects, cardiovascular defects, cleft defects, and chromosomal abnormalities.

[21] Concentration data from one month before and one month after conception were calculated for the analysis of neural tube defects, for the first two months of pregnancy for cardiac and cleft defects, and for the three months before pregnancy for chromosomal abnormalities.

A nested case-control study was carried out using data collected previously for a retrospective cohort study looking at an association between adverse pregnancy outcomes and aerial spraying of malathion in Northern California (Hertz-Picciotto et al., 1989). The cohort of women enrolled in the malathion study was identified from health records of women with medically confirmed pregnancies in 1981-1982 from three Medical Care facilities. From this cohort, cases were identified as those pregnancies that had certain adverse outcomes[22]. Controls were randomly selected from the non-cases in the cohort previously described. The non-specific agents bottled water *vs.* tap water were investigated in relation to spontaneous abortions. Specific characteristics of the tap water were not investigated in any depth, although the source of the tap water (surface water, ground water, or mixed surface and ground) was determined and linked to subjects based on the water utility serving the geographic area of the maternal residences. Information on water consumption and residential history for the period from July 1, 1981 to June 30, 1982 was solicited directly from the individual subjects chosen to take part in this study. Questionnaires were mailed to the subjects. After three mailings, phone interviews were attempted with non-respondents and those with incomplete questionnaires. The mailings and interviews took place between July 1984 and February 1985, with an 87% response rate. On the questionnaire and during the interviews the subjects were asked whether during the first three months of their pregnancies they drank bottled water either plain or in combination with beverages such as coffee, tea, orange juice, etc. If yes, how much was consumed at home and how much was consumed at work. In addition the question was asked whether the subjects drank tap water during the first three months of pregnancy and if yes, did they drink tap water more than bottled water. The questionnaires indicated that water used to make both hot and cold beverages should be included. Categories of water consumption qualitatively indicated the proportions of tap water and bottled water consumed. The categories were defined as "tap water only", "mostly tap, some bottled", "mostly bottled, some tap" and "only bottled water". This Category 2b study investigated non-specific agents using personal data obtained from individuals by interviews or questionnaires.

A retrospective cohort study from California was a re-analysis of the data of the original Deane *et al.* (Deane et al., 1989) study looking at adverse reproductive outcomes in association with exposure to contaminated ground water (Deane et al., 1992). As in the contaminated water study, the subject groups were identified based on exposure to contaminated drinking water. This exposure was determined by residence in one of two census tracts, one of which was potentially exposed to contaminated water and one of which was not. A household enumeration was carried out in these two census tracts in 1983 to determine whether any women in those households had been pregnant during 1980 or 1981, the two years during which the contamination was suspected to have started. Women who had pregnancies starting in 1980 or 1981 were interviewed for information on pregnancy outcome and water consumption behaviour. Information on water consumption included questions on the amount of cold tap water usually consumed at home during pregnancy and for the 3 months before pregnancy. If bottled water was reported as usually consumed, no information on tap water consumption was solicited. In the case of usual bottled water consumption, the amount of tap water consumed was assumed to be none for the purpose of analysis. The same questions were asked about water consumed at work. No information on water consumption was obtained for specific trimesters of pregnancy. Amount of tapwater consumed was treated as a categorical variable with the categories being (in number of glasses of cold tapwater per day): 0, 1, 2, 3, 4-6, and ≥7. No characteristics of tap water were measured in this study and specific agents were neither identified nor quantified. However, individual exposure to tap water was quantified for each subject from personal information obtained from interviews.

---

[22] Including ectopic pregnancies, spontaneous abortions, stillbirths, congenital anomalies, intrauterine growth retardation, and neonatal deaths.

A case-control study in California was one of several studies designed to clarify some unresolved issues raised by the Deane study of 1989 (Windham et al., 1992). This study was designed to investigate the possible association between tap water consumption and spontaneous abortion. In this study, cases were identified from women in Santa Clara County, California who had spontaneous abortions verified by hospital pathology laboratories and whose pregnancies had started in 1986. Two controls for each case were chosen from all live births in the County, matched to each case by calendar date of last menstrual period and hospital. This exact matching was dropped during recruitment and interviewing. The subjects were interviewed using a computer-assisted interview, which took about 40 minutes[23]. Cases were asked about events during their entire pregnancy, whereas controls were asked only about events during the first 20 weeks of pregnancy, to make the exposure periods more comparable. The source of tap water, either surface or ground, was determined and linked to individuals based on residence at pregnancy termination in a geographic area or census tract served by various water utilities. An additional analysis was done including only women who had not moved between January 1986 and the time of the interview. Analysis of risk of spontaneous abortion is presented for several different exposure variables in this study. Amount of tap water consumed at home was treated as a categorical variable with the categories being (in glasses per day) 0, 1-2, ≥3, and any (>0.5). Water type (ground, surface, or mixed) was also investigated, as was bottled water consumption. Amount of bottled water was categorized as (in glasses per day) 0, 1-2, 3-4, 5-6, 7-8, ≥9 and any (>0.5). Personal information on water consumption was linked to non-specific agents in this Category 2b study.

---

[23] Questions on the consumption of all beverages, including cold tap water, alcohol, and caffeinated beverages, were included in this interview. Each subject was asked about the consumption of beverages before pregnancy and whether consumption of any of the beverages had changed during pregnancy. If consumption had changed, subjects were asked in which week and by what amount consumption changed. Additional tap water questions included number of glasses of cold tapwater or beverages made with cold tap water usually consumed each day at home or at work. Only home tap water consumption is included in the analysis of this study to increase comparability to other studies. If half a glass of tap water or less was consumed per day, the consumption was classified as "none". Questions were also asked about bottled water and whether it was "usually" consumed at home or work during pregnancy and the amount consumed daily. Information on showering and bathing time per week was also solicited during the interviews.

Another case-control study was carried out in California as one of several studies to clarify issues raised by the Deane et al. (Deane et al., 1989) water contamination study in Santa Clara County (Fenster et al., 1992). This study was not designed to investigate an association between tap water consumption and spontaneous abortion, but rather to investigate the effect of reporting consistency on the results of studies that are designed to investigate a causal association. However, the design of this study was modeled closely after the study described immediately preceding this paragraph and can therefore be of use in investigating exposure assessment. Subjects for this case control study were recruited from two hospital facilities in Santa Clara County, California. Cases were identified as women who had a spontaneous abortion confirmed by a hospital pathology laboratory. Two controls for every case were chosen and matched to the cases by date of last menstrual period and hospital. The first interviews were conducted between June and October 1987 for women whose last menstrual periods occurred between February and July 1987, on average 24 weeks after the date of their last menstrual period. The interviews were computer-assisted telephone interviews with the same questions asked as in the Windham et al. (Windham et al., 1992) study. The second interview with both cases and controls occurred within 5 weeks of the termination of the controls' pregnancies or approximately one year after the last menstrual period of both cases and controls. Cases were asked about consumption during their entire pregnancies and exposure periods were made more comparable by interviewing controls only about events during the first 20 weeks of their pregnancies. Tap water and bottled water were each categorized by number of glasses per day into the categories (in glasses per day): 0, 1, 2, 3, 4, 5, ≥6, and any (>0.5). Odds ratios were calculated for information from both the first and second interviews. It was found that controls tended to underreport tap water consumption in the second interview compared to the first, but this effect was not seen in cases, thereby elevating the OR by 55% from the first interview to the second. This suggests the possibility of reporting bias with increasing length of time from the exposure period of interest. No such effect was observed with bottled water.

In another case-control study following up the 1989 Deane et al. study, the investigators examined a possible association between tap water consumption and various reproductive outcomes[24] in four study areas in Santa Clara County, California (Wrensch et al., 1992). The women in four study areas in this county were enumerated in 1986 to identify potential study subjects who had been pregnant sometime within the years 1980-1985. Cases included spontaneous abortions validated by physicians conceived between January 1, 1980, and March 31, 1985, and controls included live births conceived during the same time period. Cases and controls were interviewed via telephone and information on water consumption was solicited[25]. Analysis for spontaneous abortions and birth defects included water consumption information from the first trimester of pregnancy, whereas analysis for low birth weight included water consumption information from the entire pregnancy. A subsample of women was re-interviewed to obtain information on weekly duration of showering and bathing, and filter use in the home. Water consumption was broken down into several categories for the analysis. The categories for consumption of cold tap water at home (in glasses per day) were: 0, 1, and ≥2. Duration of showering and bathing per week was also divided into categories (in minutes per week): <50, 50-99, 100-149, ≥150. Water filter use was divided into filter types as follows: active and regularly serviced (including reverse osmosis, activated charcoal, and carbon filters), water softener, not-serviced/unknown type, and no water filter. To summarize, non-specific agents were investigated together with personal information on water consumption and water use behaviours.

The most recent study in Category 2b is a prospective cohort study, one of a pair of complementary studies carried out in California (Swan et al., 1998). This study investigated a possible association between the consumption of tap water and spontaneous abortion. Women in the early stages of pregnancy living in three regions throughout California were identified as potential subjects for this study. The regions were divided based on water source, classified as mixed ground and surface water, primarily surface water, or primarily ground water. Computer-assisted telephone interviews were conducted to determine water and other consumption during both the week beginning with the last menstrual period and the week before the interview[26]. The information on consumption of water was broken down in several ways for the purposes of analysis. Odds ratios were calculated for comparisons of three types of water (bottled water, total tap water and cold tap water). Bottled water comparison categories included (in glasses/day) 0.5-5.5 *vs.* 0, ≥6 vs. 0, ≥6 and no cold tap water *vs.* ≥6 and cold tap water. Both total tap water (cold tap water and hot tap water combined) and cold tap water alone were categorized as 0.5-5.5 *vs.* 0, ≥6 *vs.* 0, ≥6 and no bottled water *vs.* 0 and ≥6 bottled water. High consumption of cold tap water (meaning no bottled water intake) and high consumption of bottled water (no tap water intake) were also stratified into categories as follows: ≥2, ≥3, ≥4, ≥5, ≥6, ≥7, and ≥8. In this last Category 2b study, non-specific agents were investigated in association with personal information on water consumption and other water use behaviours.

---

[24] Spontaneous abortion, low birth weight (<2500g), and birth defects. Birth defects were validated by a review of medical records, and birth weight was obtained from birth certificates.

[25] This information included questions on the number of glasses of cold tap water per day the subjects drank at home during pregnancy or in the 3 months prior to pregnancy, if the amount differed in the first trimester, and if so how much was consumed during the first trimester. The same questions were asked about cold tap water consumption at work. Also included were questions on bottled water consumption at home during pregnancy or in the three months before pregnancy, and whether bottled water was usually consumed at home. The same questions were asked about bottled water consumption at work. Only the information on tap water consumed at home was used in the analysis reported in this study.

[26] Questions were asked regarding the consumption of four types of water: cold tap water at home and beverages made from cold tap water at home, heated tap water at home and beverages made from heated tap water at home, non-carbonated bottled water, and carbonated water (not included in the analysis). Particular notice was taken of any change in amount consumed before and after the pregnancy and if a change was noted, the time of that change was ascertained. Any treatment of tap water prior to drinking, such as letting it stand or refrigerating it, or use of water filters or

Category 3 studies

Category 3 studies combine the improvements found in both Category 2a and Category 2b studies over the Category 1 studies. In Category 3 studies, specific agents are measured and personal information is gathered from interviews with study subjects. Although concentrations of specific agents are available, these data are obtained from public records of distribution system monitoring and not directly measured in the subjects' homes.

To date there are only two studies that can be classified as Category 3 studies. One such study is the complement to the study Category 2b conducted by Swan et al. in 1998 in California (Waller et al., 1998). Since the original study was published, efforts to improve the exposure assessment have continued (Waller et al., 2001). However, the original study will be addressed here. Information from the same study subjects was for this prospective cohort study as for the previous study. The outcome of interest was spontaneous abortion. Information from the Swan et al. (Swan et al., 1998) interviews was used to determine personal water consumption and water use behaviours. The main difference between these studies, and the characteristic that classifies them into different categories, is the quantification of specific THM levels. Information on TTHM and individual THM concentrations was obtained from the monitoring data of the water utilities serving the geographic areas in which the study subjects resided. The monitoring data consisted of individual THM and TTHM concentrations from quarterly distribution system samples. TTHM levels were estimated for each subject by averaging all the distribution system TTHM measurements taken by the water utility serving the subject during the subject's first trimester. If no measurements were taken during the first trimester, then an average of measurements taken within 30 days of the subject's first trimester was used. Analogous methods were used to determine concentrations of the individual THMs for each subject's first trimester. Exposure levels were calculated using a combination of the personal information on water consumption and the concentrations estimated from monitoring data. The three levels of TTHM exposure are: zero glasses per day of

---

purifiers was also determined. Brand and company of bottled water was established. Questions regarding the duration and frequency of showering or bathing were asked, and this information was used to calculate at weekly duration of showering.

cold tap water consumed with TTHM concentrations of <75 μg/L or ≥75 μg/L; less than 5 glasses per day of cold tap water consumed with TTHM concentrations of <75 μg/L or ≥75 μg/L; five or more glasses per day or cold tap water consumed with TTHM concentrations of <75 μg/L or ≥75 μg/L. For all levels of water consumption, the referent category was the "low" exposure of <75 μg/L. This Category 3 study used specific agents and personal water consumption data to examine an association between exposure to THMs in drinking water and spontaneous abortion. However, THM concentrations were not measured directly from study subject's homes, but were obtained from water utilities.

The second study in this category was a population-based case-control study in three counties in North Carolina (Savitz et al., 1995) investigating an association between a variety of water characteristics and birth outcomes[27]. Water characteristics measured included water source, amount of water ingested, and THM concentrations. Cases in the study were identified from hospital or private clinic records. Spontaneous abortion cases were selected in Alamance County for the period from September 1988 to August 1991. Preterm deliveries and low birth weight cases were selected in Orange and Durham Counties for the period from September 1988 to August 1989 and in Alamance County for the time period from September 1988 to August 1991. Controls of live births and normal weights were matched to cases on a one-to-one ratio and identified from live deliveries immediately following a preterm or low birthweight case of the same race and hospital of delivery. Telephone interviews were conducted to determine certain water consumption factors[28]. The analysis was then restricted to subjects who were served by a public water utility and who drank one or more glasses of water per day. Subjects were linked to water utilities based on their residence in the geographical area served by a particular water utility. Quarterly average THM concentrations were obtained from the water utilities. An individual exposure "score" depended on the time period of the pregnancy identified as key for different outcomes. For miscarriage cases and controls, the fourth week of pregnancy was used to assign a THM concentration. For preterm birth and low birthweight cases and controls, the 28th week of pregnancy was used to assign a THM concentration. Based on the information gleaned from the personal interviews and the THM concentrations estimated from water utility monitoring data, several indices of exposure were identified and categorized. Water source was identified as coming from a community supply, a private well (referent category), or bottled water. Amount of water consumed per day was categorized as (in glasses of water consumed per day) 0, 1-3 (referent category), or 4+. A combined category of source x amount was stratified as: private well, 1-3 glasses per day; private well, 4+ glasses per day; community supply, 1-3 glasses per day; community supply, 4+ glasses per day; bottled water, regardless of amount. THM concentrations were categorized as: 40.8-63.3 μg/L (referent category), 63.4-82.7 μg/L, and 82.8-168.8 μg/L[29]. In addition, THM concentrations were treated as continuous variables per μg/L change for spontaneous abortion and per 50 μg/L change for the other outcomes. THM dose was calculated as glasses per day multiplied by THM concentration. The THM dose categories for spontaneous abortion are (units of μg/L x glasses per day): 40.8-139.9 (referent category); 140.0-275.0; 275.1-1171.0, with continuous treatment looking at a per 250 unit change. The TTHM dose categories for the other birth outcomes are (units of μg/L x glasses per day): 44.0-169.9 (referent category); 170.0-330.8; 330.9-1171.0, with continuous treatment looking at a per 250 unit change. This second Category 3 study used data on both personal water consumption and specific THM concentrations. However, the THM concentrations were an indirect measure obtained from water utility records.

---

[27] Birth outcomes of interest included spontaneous abortion, preterm delivery (<37 weeks gestational age) and low birth weight (<2500g).

[28] What was the primary source of drinking water in the home; was drinking water to the home supplied by a community water utility, a private well, or bottled water; approximately how many glasses of water were consumed per day around the time of the pregnancy.

Category 4 studies

Category 4 studies encompass the specific agents and personal water consumption and water use data of Category 3 studies. In addition, the specific agents in Category 4 studies are measured directly at the point of exposure for study subjects, for example, in the subject's home or place of work. There are no studies that strictly qualify as Category 4 studies. However, there is one study in which the THM concentrations were measured in the homes of a subset of subjects. Based on the majority of the study and the reported results, this study is technically placed in Category 2a and is dealt with in that category; however, it is also examined in Category 4 for illustration.

This population-based case-control study was carried out in New Jersey (Klotz and Pyrch, 1999) to investigate a possible association between exposure to DBPs and neural tube defects. Cases were identified through Birth Defects Registry records from 1993 and 1994. Approximate 10 potential controls were selected randomly for each month of 1993 and 1994 without regard for geographic location or case characteristic. Mothers of both cases and controls were contacted three months after the births. For the subsample of the population for which home tap water samples were taken, home visits occurred approximately four months after the birth, which corresponds to the critical period for neural tube development in the infants one year earlier during gestation. Tap water samples were taken during this home visit and an interview was conducted[30]. These residences were termed "index residences". Samples were analysed for THMs, total chlorine, and free chlorine. In residences obtaining water from a surface water source, samples were analysed for HAAs and HANs in the last 14 and 22 months of fieldwork, respectively. TTHM concentrations were treated categorically and separated both into 20 μg/L increments and tertiles as follows (in μg/L): increments <20 (referent category), 20-<40, 40-<60, 60-<80, 80+; tertiles <5 (referent category), 5-<40, 40+. HAN and HAA concentrations were also categorized: HANs (μg/L) <0.5 (referent category), 0.5-3.0, 3.0+; HAA (μg/L) <3 (referent category), 3-<35, 35+. Information on water consumption was used to estimate the quantity of TTHMs ingested per day and was stratified into tertiles (not stated in the published study). Part of this study satisfies Category 4 requirements by utilizing data on specific agents measured directly in subjects' homes and personal information on water use and water consumption.

---

[29] The precision of these data and in the resulting exposure categories is clearly artificial, particularly in light of the more general measures of volume of water ingested. However, these data are quoted directly from the published study.

[30] Interviews for index study subjects asked questions regarding quantities of hot and cold beverages ingested, use of tap water *vs.* bottled water, bathing and showering time, swimming pool use, and water filter use for the three months before conception and in the first trimester.

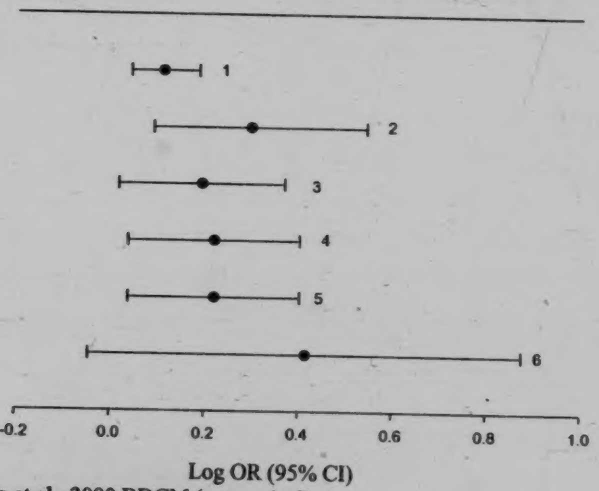
Category 5 studies

There are no Category 5 studies as the measure of exposure for Category 5 studies is a biomarker of exposure. Since appropriate biomarkers of exposure have yet to be validated, it is not possible for any Category 5 studies to have been conducted.

Study Precision

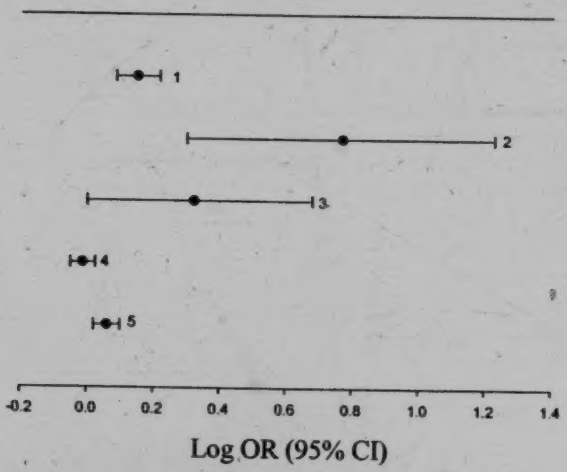
The OR and 95%CI of certain studies were plotted in Figures 2 through 5, grouped by outcome. Only outcomes examined by studies in more than one exposure category were considered. It is not possible to make any conclusions regarding the precision of studies in relation to the exposure category. The precision of studies in a function of the size and homogeneity of the sample population and not the exposure assessment. In addition, results from studies that employed multiple exposure measures are graphed separately. The individual studies are labeled by exposure category, author, and exposure measure, e.g. Category 2a: King et al., 2000 TTHM. Please note the log scale on the x-axes.

Figure 2-2. Results of epidemiology studies investigating stillbirths



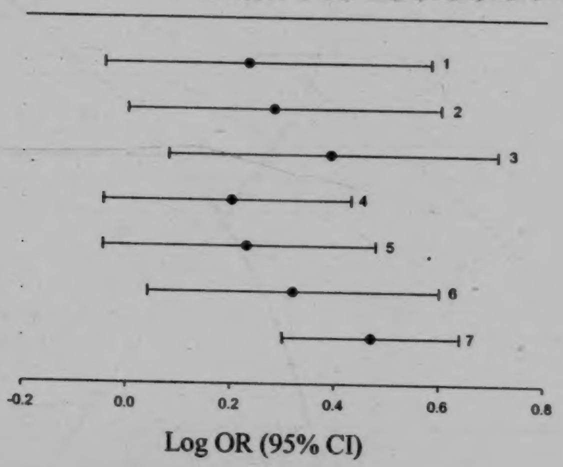
1. Category 2a: King et al., 2000 BDCM (categorical)
2. Category 2a: King et al., 2000 BDCM(continuous)
3. Category 2a: King et al., 2000 TCM
4. Category 2a: King et al., 2000 TTHM
5. Category 2a: Dodds et al., 1999
6. Category 1: Aschengrau et al., 1993 (water source and disinfection type)

Figure 2-3. Results of epidemiology studies investigating low birth weights



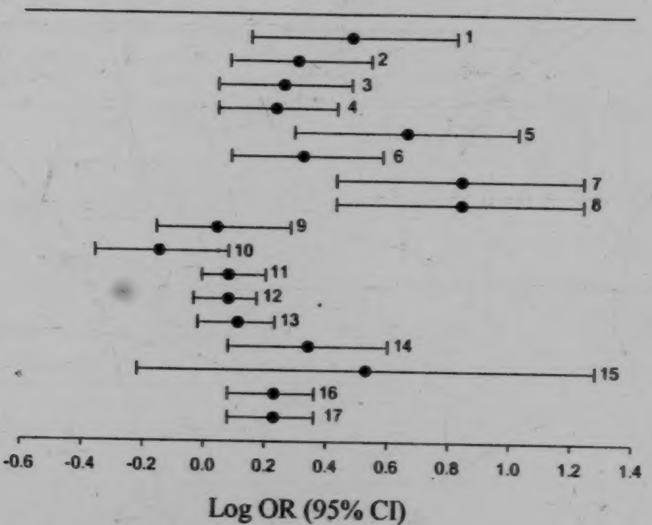
1. Category 2a: Bove et al., 1995 TTHM (50% CI)
2. Category 2a: Gallagher et al. 1998 TTHM (term low birthweight)
3. Category 2a: Gallagher et al. 1998 TTHM
4. Category 1: Jaakola et al., 2001 (measure of colour and chlorination vs. no chlorination)
5. Category 1: Kallen and Robert, 2000 (ClO_2 or NaOCl disinfection vs. no disinfection)

Figure 2-4. Results of epidemiology studies investigating neural tube defects



1. Category 4: Klotz and Pyrch, 1999 TTHM (unrestricted analysis)
2. Category 4: Klotz and Pyrch, 1999 TTHM (isolated defects)
3. Category 2a: Dodds and King, 2001 BDCM
4. Category 2a: Klotz and Pyrch, 1999 TTHM (unrestricted analysis)
5. Category 2a: Klotz and Pyrch, 1999 TTHM (isolated defects)
6. Category 2a: Klotz and Pyrch, 1999 TTHM (isolated cases, known residence)
7. Category 2a: Bove et al., 1995 TTHM (50% CI)

Figure 2-5. Results of epidemiology studies investigating spontaneous abortion



1. Category 3: Waller et al., 1998 BDCM (adjusted)
2. Category 3: Waller et al., 1998 BDCM (not adjusted)
3. Category 3: Waller et al., 1998 TTHM
4. Category 3: Savitz et al., 1995 TTHM
5. Category 2b: Swan et al., 1998 (high tap and no bottled vs. high bottled and no tap)
6. Category 2b: Swan et al., 1998 (6+ glasses tap/day vs. no tap)
7. Category 2b: Wrensch et al., 1992 (any tap vs. no tap)
8. Category 2b: Fenster et al., 1992 (any tap vs. no tap 2nd interview)
9. Category 2b: Fenster et al., 1992 (any tap vs. no tap 1st interview)
10. Category 2b: Windham et al., 1992 (0.5+ glasses/day vs. none)
11. Category 2b: Windham et al., 1992 (3+ glasses/day vs. none)
12. Category 2b: Windham et al., 1992 (1-2 glasses/day vs. none)
13. Category 2b: Deane et al., 1992 (tap water consumption)
14. Category 2b: Deane et al., 1992 (tap controlled for bottled)
15. Category 2b: Hertz-Picciotto et al., 1992 (any tap vs. only bottled)

Discussion

Having grouped the existing studies into categories, we can now endeavor to answer the questions asked at the beginning of this paper:

-How well have reproductive epidemiology studies to date characterized individual exposure to possible causal agents in treated drinking water and have there been useful improvements in exposure assessment over time?

-Can improved exposure assessment clarify a possible causal association between exposure to agents in treated drinking water and adverse reproductive effects?

-How can future epidemiology studies improve exposure assessment in the absence of biomarkers of exposure?

In this framework, there is an assumption that the categories represent a trend from the least accurate to the most accurate individual exposure assessment. There is also a clear boundary between Categories 1, 2a, 2b, and Categories 3, 4, and 5. The first three categories lack one or both of the key elements of specific agents and individual data. Categories 3, 4, and 5 include both specific agents and individual data to varying degrees of detail and methods of measurement.

Examining the studies as they are categorized in this framework, it is clear that most of the studies (20 of 23) are found in Categories 1, 2a, or 2b. Only 3 of 23 studies are in Categories 3 or 4 and no studies have Category 5 exposure assessment. In addition, the sole Category 4 study is in fact a sub-element of a Category 2a study. To further break it down, there are seven Category 1 studies, seven Category 2a studies, six Category 2b studies, two Category 3 studies, and one Category 4 study. This confirms the fact mentioned in previously published articles (Nieuwenhuijsen et al., 2000a; Nieuwenhuijsen et al., 2000b; Swan and Waller, 1998; Reif et al., 1996) that individual exposure assessment has not been well characterized to date.

This framework also allows an examination of any improvements over time; in other words, whether exposure assessment has improved in more recent studies compared to prior studies. The answer is both yes and no. The very first study, published in 1982 (Tuthill et al., 1982), was a Category 1 study and employed non-specific agents and aggregated population data. The Category 3 and 4 studies were published in the mid- to late 1990's (Savitz et al., 1995; Waller et al., 1998; Klotz and Pyrch, 1999) and characterized individual exposure assessment using specific agents and individual data. However, the most recent studies are classified as either Category 1 (Jaakola et al., 2001; Kallen and Robert, 2000; Yang et al., 2000; Magnus et al., 1999) or Category 2a (Dodds and King, 2001; King et al., 2000; Dodds et al., 1999; Klotz and Pyrch, 1999). Category 2a studies assess individual exposure using specific agents and aggregated population data. Again, the Klotz and Pyrch study (Klotz and Pyrch, 1999) is classified as both Category 2a and Category 4, with the main part of the study employing the specific agents and aggregated population data of Category 2a, while a sub-section of the study employed directly measured specific agents and individual data allowing it to be classified in Category 4.

In order to assess a causal association between an exposure and an outcome, it is helpful to look to the criteria for causality. By examining the studies as classified into various categories according to their contributions to the causality criteria, it is possible to determine in part the contribution of individual exposure assessment to the investigation of a causal association. Most of the criteria are discussed here, with the exception of biological plausibility and reversibility. Biological plausibility is best investigated using the toxicology studies mentioned previously while reversibility can not apply to these studies due to the outcomes of interest[31].

Temporality: the cause must precede the effect. All the studies were careful to define their exposure periods as occurring before the outcomes. The studies using aggregated population data (Categories 1 and 2a) generally assumed that the maternal residence at pregnancy outcome was the maternal residence for the entire pregnancy, unless information was provided otherwise in public records. In making this assumption, the additional assumption was made that there was no change in exposure levels (type of disinfection or water source in Category 1 or specific agent concentrations in Category 2a) either from inherent temporal and spatial variations in the case of specific agents or from maternal migration during pregnancy. Exposure in Category 2b studies was defined by personal data on consumption of tap water or bottled water (non-specific agents). Personal data were generally collected for the periods before and during pregnancy and therefore fulfill the temporality criterion. The issue of maternal migration does not affect these studies due to the non-specific nature of the causal agents. Category 3 studies utilized personal data to determine the amount of water consumed before and during pregnancy, as well as information on maternal residence during pregnancy. However, the concentrations of the specific agents were determined by indirect measurement and therefore are subject to temporal and spatial variations between the sampling point and the point of individual exposure. Certain assumptions about the consistency of specific agent concentrations were made in these studies in order to satisfy the temporality criterion. The sub-study of the Klotz and Pyrch (1999) study classified as a Category 4 study utilized direct measurement of specific agents at the point of exposure as well as personal data on water use. Again, an assumption was made of temporal consistency in specific agent concentration as samples were taken at the point of exposure one year after the exposure time period of interest in order to satisfy the temporality criterion. All the studies were careful to satisfy the temporality criterion, some with stated assumptions regarding maternal residence or specific agent concentrations. However, those studies employing personal data on water use and maternal residence have the advantage in determining a temporal relationship between exposure and outcome.

---

[31] A birth defect or other adverse reproductive outcome that has already occurred will not cease to exist if exposure to chlorinated water is terminated.

Consistency: if several different studies with a variety of designs, carried out in different locations and under different conditions, consistently report the same result then the argument for causality is strengthened. The criterion of consistency is difficult to fulfil in these studies because of the many differences in outcome of interest and type of exposure. There are too few studies looking at similar exposures and outcomes to make a real comparison. In addition, there are too few studies of varying design at the higher exposure categories to really examine this criterion in terms of the effect of improving individual exposure assessment.

In addition, the range of possible causal agents investigated in the studies, even in studies within the same category, contributes to the problem of too many types and levels of exposure. Table 2-2 summarizes the agents of exposure for each study.

Table 2-2. Summary of Agents of Exposure

Category	Study	Suspected Causal Agent	Exposure Range	Referent Level
1	Tuthill et al., 1982	Chlorine dioxide	Chlorine dioxide: yes/no Chlorine: yes/no	Chlorinated water (chlorine)
	Aschengrau et al., 1993	Chlorination (chlorine)	Chlorine: yes/no Chloramine: yes/no	Chloraminated water
	Kanitz et al., 1996	Chlorine dioxide, sodium hypochlorite, both	Chlorine dioxide: yes/no Sodium hypochlorite: yes/no Both: yes/no	No disinfection
	Magnus et al., 1999	Chlorination (chlorine) and colour	Chlorine: high/low Colour: high/low	Low chlorine and low colour
	Yang et al., 2000	Chlorination (chlorine)	Chlorine: yes/no	No disinfection
	Kallen and Robert, 2000	Chlorine dioxide, sodium hypochlorite	Chlorine dioxide: yes/no Sodium hypochlorite: yes/no	No disinfection
	Jaakola et al., 2001	Chlorination (chlorine) and colour	Chlorine: high/low Colour: high/low	Low chlorine and low colour
2a	Kramer et al., 1992	TCM, TBM, BDCM, CDBM	TCM and BDCM: undetectable (<1μg/L), low (1-9μg/L), high (≥10μg/L) CDBM and TBM: undetectable (<1μg/L), low (1-3μg/L), high (≥4μg/L)	All compounds: undetectable (<1μg/L)

Table 2-2, continued. Summary of Agents of Exposure

Category	Study	Suspected Causal Agent	Exposure Range	Referent Level
2a continued	Bove et al., 1995	TTHM	≤20ppb,>20-40 ppb, >40-60 ppb, >60-80 ppb, >80-100 ppb, >100 ppb	≤20ppb
	Gallagher et al., 1998	TTHM	≤20ppb, 21-40ppb, 41-60ppb, ≥61ppb	≤20ppb
	Klotz and Pyrch, 1999	TTHM	Increments:<20ppb, 20-<40ppb, 40-<60ppb, 60-<80ppb, 80+ppb Tertiles: <5ppb, 5-40ppb, 40+ppb	Increments: <20ppb Tertiles: <5ppb
	Dodds et al., 1999	TTHM	0-49μg/L, 50-74μg/L, 75-99μg/L, ≥100μg/L	0-49μg/L
	King et al., 2000	TTHM, TCM, BDCM	TTHM and TCM: <50 μg/L, 50-74μg/L, 75-99μg/L, ≥100μg/L; per 10μg/L BDCM: <5μg/L, 5-9μg/L, 10-19μg/L, ≥20μg/L; per 10μg/L	TTHM and TCM: <50μg/L BDCM: <5μg/L
	Dodds and King, 2001	TCM, BDCM	TCM: <50 μg/L, 50-74μg/L, 75-99μg/L, ≥100μg/L BDCM: <5μg/L, 5-9μg/L, 10-19μg/L, ≥20μg/L	TCM: <50μg/L BDCM: <5μg/L

Table 2-2, continued. Summary of Agents of Exposure

Category	Study	Suspected Causal Agent	Exposure Range	Referent Level
2b	Hertz-Picciotto et al., 1989	Chlorinated tap water	Chlorinated tap water: mostly, only Bottled water: mostly, only	Only or mostly bottled water; only bottled water
	Deane et al., 1992	Chlorinated tap water	Number of glasses cold home tap water/day: 0, 1, 2, 3, 4-6, ≥7; Also: 1, 1-3, ≥4	0 glasses cold home tap water/day
	Windham et al., 1992	Chlorinated tap water	Cold home tap water in glasses/day: 0, 1-2, ≥3; any (≥0.5)	0 glasses cold home tap water/day
	Fenster et al., 1992	Chlorinated tap water	Cold home tap water in glasses/day: 0, 1, 2, 3, 4, 5, ≥6; any (≥0.5)	0 glasses cold home tap water/day
	Wrensch et al., 1992	Chlorinated tap water	Cold home tap water in glasses/day: 0, 1, ≥2	0 glasses cold home tap water/day
	Swan et al., 1998	Chlorinated tap water	Cold home tap water in glasses/day: 0, 0.5-5.5, ≥6	0 glasses cold home tap water/day

Table 2-2, continued. Summary of Agents of Exposure

Category	Study	Suspected Causal Agent	Exposure Range	Referent Level
3	Savitz et al., 1995	TTHM	TTHM dose (TTHM conc. in µg/L x glasses/day): 40.8-139.9, 140.0-275.0, 275.1-1171.0; per 250 unit change[a]	40.8-139.9 (TTHM conc. in µg/L x glasses/day):
	Waller et al., 1998	TTHM, BDCM, TCM, TBM, BDCM, CDBM	Glasses cold home tap water/day: 0, <5, ≥5 TTHM: <75µg/L, ≥75µg/L TCM: <17µg/L, ≥17µg/L TBM: <16µg/L, ≥16µg/L BDCM: <18µg/L, ≥18µg/L CDBM: <31µg/L, ≥31µg/L	< 5 glasses of cold home water/day and <75µg/L TTHM, or <17µg/L TCM, or <16µg/L TBM, or <18µg/L BDCM, or <31µg/L CDBM
4	Klotz and Pyrch, 1999	TTHM	Increments:<20ppb, 20-<40ppb, 40-<60ppb, 60-<80ppb, 80+ppb Tertiles: <5ppb, 5-40ppb, 40+ppb	Increments: <20ppb Tertiles: <5ppb

[a] The precision of these data is likely unwarranted. The data were taken directly from the published study.

Within Category 1, non-specific agents range from simply chlorinated vs. non-chlorinated water to the more complicated chlorination/no chlorination coupled with a measure of colour. In addition, the referent levels vary from study to study. In some cases in Category 1, the referent level of one study is the suspected causal agent of another. In Category 2b, where non-specific agents are coupled with personal ingestion exposure, the suspected causal agent (chlorinated tap water) is consistent. However, the increments of exposure garnered from personal information vary. Most Category 2b studies use increments of "glasses of cold home tap water per day" as a means of categorizing exposure levels. However, the cut-off points for the exposure levels in each study vary substantially. In two of the six studies (Fenster, 1992; Swan, 1998), the highest exposure level is ≥6 glasses/day and one study gives ≥7 glasses/day (Deane et al., 1992), whereas in another study (Wrensch et al., 1992) the highest exposure level is ≥2 glasses/day.

There is consistent use of zero glasses of cold home tap water per day as a referent level in Category 2b. A broad comparison of Category 1 and Category 2b studies suggests there is more consistency in Category 2b than in Category 1 in the agent under investigation as well as the referent level. Whether this is a result of improved exposure assessment or the circumstances provoking the initiation of the Category 2b studies is a matter of question. Improved exposure assessment in these studies can provide the opportunity for greater comparison between studies only if both exposure levels and referent levels are consistent.

In Categories 2a, 3, and 4, actual compounds are identified as potential causal agents, which include TTHMs and individual THMs. However, the exposure levels used in these Categories exhibit little consistency. For those studies investigating only TTHMs (Bove et al, 1995; Gallagher et al., 1998; Klotz and Pyrch, 1999; Dodds et al., 1999; King et al., 2000; Savitz et al., 1995; Waller et al., 1998), the exposures levels are developed from various increments of TTHM concentration. Four of the studies divide TTHM concentration into 20 μg/L increments (Bove, 1995; Gallagher et al., 1998; Klotz and Pyrch, 1999[32]), two divide THM concentration into 50 μg/L increments (Dodds et al., 1999; King et al., 2000), and one study gives a dose related to TTHM concentrations in non-standardized increments (Savitz et al., 1995). The Klotz and Pyrch (1999) studies also divide TTHM concentrations into tertiles of non-equal increments. The reference levels in the studies looking at TTHM exposure vary greatly. In those looking at TTHM concentration and aggregated population data the referent levels vary from <5ppb to <50ppb (Bove et al., 1995; et al., 1998; Klotz and Pyrch, 1999; Dodds et al., 1999; King et al., 2000). The two studies with TTHM exposure and personal consumption information (Savitz et al., 1995; Waller et al., 1998) each use different methods of determining exposure, resulting in different exposure and referent levels that are not easily compared. Both studies determine the number of glasses consumed per day and the THM concentrations. However, the Savitz et al. study combines these two pieces of information to determine a dose (THM concentration X glasses of water consumed per day). The Waller et al. study does not make this calculation and the resulting exposure levels are based on combinations of these variables. The Category 4 study uses only TTHM concentrations and does not attempt to calculate a dose using the personal information obtained in the study.

In addition, some studies in these categories investigated individual THMs as potential causal agents (Kramer et al., 1992; King et al., 2000; Dodds and King, 2001; Waller et al., 1998). Again, the exposure levels and referent levels among these studies show great variability. In particular, the levels for Kramer et al. (1992) can be compared to those of King et al. (2000). Both studies investigated TCM and BDCM; however, the exposure levels used are quite different. The TCM levels in the Kramer et al (1992) study range from ≤1μg/L (undetectable, the lowest level) to ≥10 μg/L (the highest level). The TCM levels in the King et al. (2000) study range from <50μg/L (the lowest level) to ≥100μg/L (the highest level). Both of these studies reported statistically significant ORs at their highest levels of exposure. While this comparison may be considered to be unfair given the different outcomes investigated by these studies and perhaps other water quality factors between the study locations, it is interesting to note that the highest exposure level in one study is well within the referent level of the other, and *they still both found statistically significant results*.

---

[32] Klotz and Pyrch is counted twice, being in both Categories 2a and 4.

The issue of exposure levels outlines a problem underlying the epidemiology studies on this topic to date. It is difficult to determine whether there is consensus among studies when the exposure levels and referent levels used in the studies are not comparable. This begs the question whether similar results would be found between the studies and even within each study if similar exposure levels and referent levels were used between all the studies within each category. In dividing the studies into categories based on their exposure assessment methods, this framework provides a tool for identifying specific problems in the body of work to date and also for identifying specific improvements for future studies. In this case, a suggestion for future studies would be to select exposure levels and referent levels to allow meaningful comparisons with previous studies as much as possible. The objective should be to provide a coherent, comparable body of work from which the question of whether or not a relationship exists between adverse reproductive outcomes and specific agents in chlorinated drinking water can be answered.

There is a potential problem unique to studies investigating exposure to bottled water vs. tap water. In general, the population of people who make the lifestyle choice to drink bottled water and can afford to drink primarily bottled water is different from the population of people who either choose not to drink bottled water primarily or can not afford to. This is a possible source of confounding that is difficult to account for. The populations being compared may be very different from one another in aspects that may affect reproductive outcomes.

Dose-Response Relationship: If the frequency or severity of the outcome increases with increasing frequency or magnitude of exposure to the potential causal agent, then a dose-response relationship is seen. Dose-response relationships can be seen only in Categories 2a and higher as Category 1 studies have no measure of dose. A dose-response relationship was seen only in one study in each of Categories 2a and 2b. No dose-response relationship was seen in the Category 3 or 4 studies. It is difficult to compare these two studies because of their different exposure assessment methods, exposure levels, and outcomes, as well as the limited number of studies at the higher exposure categories. This does re-emphasize the need for quantifiable exposure assessment as well as exposure assessment that is consistent and comparable between studies.

Strength of Association: measured by the odds ratio (OR) or relative risk (RR) of the study. A large OR or RR argues more strongly for causality than does a small one. Again, it is difficult to make any strong conclusions based on exposure assessment because of the small number of studies and the different exposures and outcomes investigated. However, there are studies in different categories that are similar enough to provide some comparison with regards to this criterion. Two such studies are the Klotz and Pyrch (1999) Category 2a study and the Klotz and Pyrch (1999) Category 4 sub-study. These studies are essentially the same study, looking at the same outcomes with the same study subjects and the same specific agents. The differences lie in the measurement of the specific agents (directly or indirectly) and the use of personal *vs.* aggregated population data. Comparing the results from TTHMs at 40+ ppb (highest tertile *vs.* lowest tertile), there is a slight difference in the ORs between the two studies. The Category 2a study shows an unrestricted OR of 1.6 (95%CI: 0.9-2.70) whereas the Category 4 sub-study shows the same unrestricted OR as 1.7 (95%CI: 0.9-3.8). When the analysis was restricted to isolated defects, the Category 2a study gave an OR of 1.7 (95%CI: 0.9-3.0) and the Category 4 sub-study gave an OR of 1.9 (95%CI 1.0-4.0). In both the restricted and the unrestricted analysis the Category 4 sub-study with better individual exposure assessment gave slightly higher ORs than the Category 2a study. This is only one comparison and therefore the suggestion that better exposure assessment will certainly result in stronger associations cannot be confirmed, but this one observation might be interpreted as at least being encouraging.

Although it is impossible to make meaningful, broad comparisons across the studies, a look at the median ORs and range of ORs for the studies grouped in different ways can be useful. Table 2-3 is a summary of the ranges of and median ORs for various groupings.

Table 2-3. Summary of OR medians and ranges[a]

Grouping	Number of studies	OR median	OR range
Category 1	7	2.0	1.09-3.5
Category 2a	7	1.9	1.29-5.9
Category 2b	6	2.2	1.2-6.9
Category 3	2	2.4	1.8-3.0
Category 4	1	N/A	N/A
Non-specific agents (Categories 1, 2b)	13	2.0	1.09-6.9
Specific agents (Categories 2a, 3, 4)	10	1.9	1.29-5.9
Aggregated population data (Categories 1, 2a)	14	1.9	1.09-5.9
Personal data (Categories 2b, 3, 4)	9	2.0	1.2-6.9

[a] The values presented in this table are based solely on the statistically significant OR results presented in Table 2-1. Great care must be taken when interpreting these values.

Although great care must be taken when interpreting the values in Table 2-3, there are several points of interest. The first is that there seems to be a general, although slight, increase in the median OR from the lower Categories to the higher Categories. This could encourage a belief that better exposure assessment over-all will give a better indication of a causal relationship. To further investigate this potential finding, the studies were further analyzed according to their use of specific *vs.* non-specific agents and aggregated population data *vs.* personal data. If we are to make a sweeping statement interpreting these data as indicating that better exposure assessment gives more accurate ORs, we would expect that the median ORs for the specific agents would be larger than the median ORs for the non-specific agents. This is based on the assumption that there is indeed a link between specific agents in drinking water and the reproductive outcomes examined in these studies. The same logic would apply to aggregated population *vs.* personal data. The summary in Table 2-3 does not support this logic. The table shows that the median ORs for both sets of groupings is approximately 2, and there is virtually no difference in the median ORs upon improved exposure assessment. In the case of specific or non-specific agents, this could indicate that the wrong agents are being measured, and the agents that are being measured are not good surrogates for the actual causal agents. Alternately, the measures used to quantify the specific agents may not be giving an accurate indication of concentrations to which the study subjects are exposed. In the case of the aggregated vs. population data, the lack of improvement in the median OR could be due to factors affecting the subjects' exposures to DBPs in chlorinated drinking water. Most of the studies employing personal exposure concentrated mainly on ingestion of cold tap water, which has been shown not to provide the whole picture of exposure to DBPs in treated water (Lin and Hoang, 2000; Weisel and

Jo, 1996; Levesque et al., 1994; Jo et al., 1990). In addition, significant intra-individual variability in daily ingestion of raw water has been found in recent studies (Bader et al., Submitted; Froese et al., 2002, In Press) which could contribute to inaccurate personal ingestion data collected via interviews and questionnaires.

A look at the ranges of the ORs in addition to the median ORs for the various groups may prove to be a useful exercise. Ideally, improved exposure assessment would result in smaller ranges of ORs as studies converge on the "true" OR. This pattern is neither seen going from low to high Categories, nor from aggregated to personal data. A decrease in the OR range is seen between the studies employing non-specific data and those using specific data; however, the difference is very slight.

Grouping the studies according to their exposure assessment methods gives an approach for comparing the studies, identifying weaknesses, and providing suggestions for improvement of future studies. Using a summary of the medians and ranges of ORs, we can cautiously postulate that improving exposure assessment in epidemiology studies may result in more accurate ORs. However, this idea is not supported by an examination of the exposure assessment elements separately. It may be that there are so many factors contributing to variations in the exposure data that small one- or two-Category improvements in exposure assessment, especially in the lower Categories, cannot overcome the noise in the resulting strengths of association. It is possible that more improved exposure assessment in the higher Categories (Categories 3, 4, and 5) will eliminate some of this noise. However, there are so few studies in these Categories (and none in Category 5) that it is difficult to predict the effect of better exposure assessment. Again, a suggestion for future studies is to investigate directly-measured, specific agents and personal exposure information, at least for sub-sections of larger studies.

It should be noted that the assumption that better exposure assessment will result in stronger measures of association only applies if there really is in fact a causal association. If there is no causal association, better and unbiased individual exposure assessment will reflect that fact.

Study design: There are many different epidemiologic study designs, each with a different level of ability to test for causation. The best study design for testing causation is the randomized controlled double-blind trial or clinical trial, followed by cohort studies, case-control studies, and cross-sectional studies. The majority of the studies in both the upper and lower categories are case-control or cohort studies. There are a few cross-sectional studies in Categories 1 and 2a, and an aggregated study in Category 1. In general, the higher exposure categories employ stronger study designs, such as case-control or cohort studies, as a result of the more detailed exposure assessment. Some of the studies that are considered case-control (Aschengrau et al., 1993; Kramer et al., 1992; Klotz and Pyrch, 1999

(Category 2a-section)) or cohort (Magnus et al., 1999; Yang et al., 2000; Kallen and Robert, 2000; Jaakola et al., 2001;Gallagher et al., 1998; Dodds et al., 1999; King et al., 2000; Dodds and King, 2001) have not used individual exposure assessment.

In the absence of biomarkers of exposure, much can be done to improve exposure assessment in future epidemiology studies. From the categorical classification of the epidemiology studies using the framework introduced in this paper and the paucity of studies in the greater quality exposure categories, it can be seen that much must be done to improve individual exposure assessment. In particular, future studies must concentrate on measuring specific causal agents, including THMs, but also other DBPs in the water suggested as possible causal agents by toxicology studies, such as HAAs and HANs. Concentrations of these agents should be measured at the points of exposure. This includes taking into account maternal migration during pregnancy, exposure outside the home, time of exposure relative to the outcome of interest, and water use activities. Personal data on water consumption and water use activities is also necessary to give a complete assessment of exposure. Granted, this "complete" exposure assessment is costly and time-consuming. However, the studies themselves are costly and we must ask whether there is any utility in performing more studies with the same limitations that are apparent in those studies done to date. In addition, many questions need to be raised as to the best methods of performing this "complete" exposure assessment. However, it should be possible to incorporate this complete exposure assessment into sub-sections of future epidemiology studies, thereby providing a means of comparing the utility of improving exposure assessment over present levels, and ultimately, leading to the answer sought by the epidemiology studies: is there indeed a causal association between agents in treated drinking water and adverse reproductive outcomes?

Conclusions

In this Chapter the problem of individual exposure assessment was outlined for epidemiology studies investigating a possible causal association between agents in chlorinated drinking water and adverse reproductive effects. In short, problems in individual exposure assessment stem from variations in disinfection by-product concentrations from formation in the water treatment plants to the points of exposure, differences in exposure routes and routes of uptake of DBPs into the body, and a lack of knowledge as to the specific causal agent or agents. Biomarkers of exposure have been suggested to be the "gold standard" measure for determining exposure. However, it is unlikely that an ideal biomarker of exposure will be validated for use in the immediate future. In the mean time, exposure assessment in epidemiology studies must utilize established methods.

Several questions were introduced regarding the individual exposure assessment employed in epidemiology studies to date, as well as improvements for future studies. A framework was introduced to aid in the comparison of epidemiology studies by categorizing the studies in terms of the exposure assessment employed.

Exposure assessment was broken down into the key elements of non-specific or specific agents, measured directly or indirectly, and personal or aggregated population data. The framework showed that very few studies have employed higher level exposure assessment (Categories 3 and 4), investigating both specific agents and personal data. In addition, there does not seem to be an observable improvement in exposure assessment over time, with the most recent studies classified in lower Categories 1 and 2a. The framework also allowed the identification of some weaknesses of the studies as a body of work. One glaring weakness is the lack of consistent exposure levels and referent levels among studies looking at the same exposure elements (specific vs. nonspecific agents; personal vs. aggregated data). This is an important issue to address in future studies, particularly those employing more sophisticated exposure assessment. Investigators need to plan their studies in light of the evidence available from prior studies, including the measures used for individual exposure assessment. In addition, analysis with the framework suggested that the noise associated with less sophisticated methods of exposure assessment can overcome the minor improvements in exposure assessment between Categories 1, 2a, and 2b. Hence, it is even more important for future epidemiology studies to employ exposure assessment from Categories 3 and 4. There is little merit in having more studies performed with inadequate exposure assessment. Ultimately, biomarkers of exposure will be validated and studies will be done with Category 5-level exposure assessment.

The framework presented in this section provides an approach for investigating the effect of improvements in exposure assessment on the outcome of epidemiology studies. Unfortunately, studies to date have investigated very different exposures and outcomes, making it difficult to compare among studies. The paucity of studies at the higher levels adds to the difficulty of establishing meaningful comparisons. In addition, the framework provides an approach to identifying weaknesses in the studies and suggestions for improvements in future studies. As studies to date have been categorically shown to have employed less than optimal exposure assessment, it is suggested that future studies concentrate on specific agents and personal data on water consumption and water use activities.

References

Aschengrau, A., S. Zierler and A. Cohen, 1993. Quality of Community Drinking Water and the Occurrence of Late Adverse Pregnancy Outcomes. *Archives of Environmental Health* **48**(2): 105-113.

Bader, Erin, Kenneth L. Froese and Steve E. Hrudey, Submitted. Determination of Elimination Half-Life for TCAA in a Controlled Human Exposure Study. .

Bove, Frank J., Mark C. Fulcome, Judith B. Klotz, Jorge Esmart, Ellen M Dufficy and Jonathan E. Savrin, 1995. Public Drinking Water Contamination and Birth Outcomes. *American Journal of Epidemiology* **141**(9): 850-862.

Deane, Margaret, Shanna H. Swan, John A. Harris, David M. Epstein and Raymond R. Neutra, 1989. Adverse Pregnancy Outcomes in Relation to Water Contamination, Santa Clara County, California, 1980-1981. *American Journal of Epidemiology* **129**(5): 894-904.

Deane, Margaret, Shanna H. Swan, John A. Harris, David M. Epstein and Raymond R. Neutra, 1992. Adverse Pregnancy Outcomes in Relation to Water Consumption: A Re-analysis of Data from the Original Santa Clara County Study, California 1980-1981. *Epidemiology* **3**(2): 94-97.

Dodds, Linda and Will King, 2001. Relation Between Trihalomethane Compounds and Birth Defects. *Occupational and Environmental Medicine* **58**: 443-446.

Dodds, Linda, Will King, Christy Woolcott and Jason Pole, 1999. Trihalomethanes in Public Water Supplies and Adverse Birth Outcomes. *Epidemiology* **10**(3): 233-237.

Fenster, Laura, Gayle C. Windham, Shanna H. Swan, David M. Epstein and Raymond R. Neutra, 1992. Tap or Bottled Water Consumption and Spontaneous Abortion in a Case-Control Study of Reporting Consistency. *Epidemiology* **3**(2): 120-124.

Froese, Kenneth L., Martha I. Sinclair and Steve E. Hrudey, 2002, In Press. Trichloroacetic Acid as a Biomarker of Human Exposure to DBPs in Drinking Water - Adelaide Trial. *Environmental Health Perspectives* **110**.

Gallagher, Michael D., John R. Nuckols, Lorann Stallones and David A. Savitz, 1998. Exposure to Trihalomethanes and Adverse Pregnancy Outcomes. *Epidemiology* **9**(5): 484-489.

Hertz-Picciotto, Irva, Shanna H. Swan, Raymond R. Neutra and Steven J. Samuels, 1989. Spontaneous Abortions in Relation to Consumption of Tap Water:

An Application of Methods from Survival Analysis to a Pregnancy Follow-Up Study. *American Journal of epidemiology* **130**(1): 79-93.

Jaakola, J.J.K., P. Magnus, A. Skrondal, B.-F. Hwang, G. Becher and E. Dybing, 2001. Foetal Growth and Duration of Gestation Relative to Water Chlorination. *Occupational and Environmental Medicine* **58**: 437-442.

Jo, Wan K., Clifford P. Weisel and Paul J. Lioy, 1990. Routes of Chloroform Exposure and Body Burden from Showering with Chlorinated Tap Water. *Risk Analysis* **10**(4): 575-585.

Kallen, Bengt A.J. and Elisabeth Robert, 2000. Drinking Water chlorination and Delivery Outcome--A Registry-Based Study in Sweden. *Reproductive Toxicology* **14**: 303-309.

Kanitz, Stefano , Yaurana Franco, Vittoriana Patrone, Marta Caltabellotta, Enrico Raffo, Cristina Riggiet al., 1996. Association Between Drinking Water Disinfection and Somatic Parameters at Birth. *Environmental Health Perspectives* **104**(5): 516-520.

King, Will D., Linda Dodds and Alexander C. Allen, 2000. Relation Between Stillbirth and Specific Chlorination By-Products in Public Water Supplies. *Environmental Health Perspectives* **108**(9): 883-886.

Klotz, Judith B. and Laurie A. Pyrch, 1999. Neural Tube Defects and Drinking Water Disinfection By-Products. *Epidemiology* **10**(4): 383-390.

Kramer, Michael D. , Charles F. Lynch, Peter Isacson and James W. Hanson, 1992. The Association of Waterborne Chloroform with Intrauterine Growth Retardation. *Epidemiology* **3**(4): 407-413.

Levesque, Benoit, Pierre Ayotte, Alain LeBlanc, Eric Dewailly, Denis Prud'Homme, Robert Lavoie et al., 1994. Evaluation of Dermal Respiratory Chloroform Exposure in Humans. *Environmental Health Perspectives* **102**(12): 1082-1087.

Lin, Tsair-Fuh and Shih-Wen Hoang, 2000. Inhalation Exposure to THMs from Drinking Water in South Taiwan. *The Science of the Total Environment* **246**: 41-49.

Magnus, Per, Jouni J.K. Jaakkola, Anders Skrondal, Jan Alexander, Georg Becher, Truls Kroghet al., 1999. Water Chlorination and Birth Defects. *Epidemiology* **10**(5): 513-517.

Nieuwenhuijsen, Mark J., Mireille B. Toledano, Naomi E. Eaton, John Fawell and Paul Elliott, 2000a. Chlorination Disinfection By-Products in Water and Their

Association with Adverse Reproductive Outcomes: A Review. *Occupational and Environmental Medicine* **57**: 73-85.

Nieuwenhuijsen, Mark J., Mireille B. Toledano and Paul Elliott, 2000b. Uptake of Chlorination Disinfection By-Products; A Review and a Discussion of its Implications for Exposure Assessment in Epidemiological Studies. *Journal of Exposure Analysis and Environmental Epidemiology* **10**: 586-599.

Reif, John S. , Maureen C. Hatch, Michael Bracken, Lewis B. Holmes, Bernard A. Schwetz and Philip C. Singer, 1996. Reproductive and Developmental Effects of Disinfection By-Products in Drinking Water. *Environmental Health Perspectives* **104**(10): 1056-1061.

Savitz, David A., Kurtis W. Andrews and Lisa M. Pastore, 1995. Drinking Water and Pregnancy Outcome in Central North Carolina: Source, Amount, and Trihalomethane Levels. *Environmental Health Perspectives* **103**(6): 592-596.

Swan, Shanna H. and Kirsten Waller, 1998. Disinfection By-Products and Adverse Pregnancy Outcomes: What Is the Agent and How Should it Be Measured? *Epidemiology* **9**(5): 479-481.

Swan, Shanna H., Kirsten Waller, Barbara Hopkins, Gayle Windham, Laura Fenster, Catherine Schaeferet al., 1998. A Prospective Study of Spontaneous Abortion: Relation to Amount and Source of Drinking Water Consumed in Early Pregnancy. *Epidemiology* **9**(2): 126-133.

Tuthill, R.W., R.A. Giusti, G.S. Moore and E.J. Calabrese, 1982. Health Effects Among Newborns after Prenatal Exposure to ClO_2-Disinfected Drinking Water. *Environmental Health Perspectives* **46**: 39-45.

Waller, Kirsten , Shanna H. Swan, Gerald DeLorenze and Barbara Hopkins, 1998. Trihalomethanes in Drinking Water and Spontaneous Abortion. *Epidemiology* **9**(2): 134-140.

Waller, Kirsten, Shanna H. Swan, Gayle C. Windham and Laura Fenster, 2001. Influence of Exposure Assessment Methods on Risk Estimates in an Epidemiologic Study of Total Trihalomethane Exposure and Spontaneous Abortion. *Journal of Exposure Analysis and Environmental Epidemiology* **11**: 522-531.

Weisel, Clifford and Wan-Kuen Jo, 1996. Ingestion, Inhalation, and Dermal Exposures to Chloroform and Trichloroethene from Tap Water. *Environmental Health Perspectives* **104**(1): 48-51.

Windham, Gayle C., Shanna H. Swan, Laura Fenster and Raymond R. Neutra, 1992. Tap or Bottled Water Consumption and Spontaneous Abortion: A 1986 Case-Control Study in California. *Epidemiology* 3(2): 113-119.

Wrensch, Margaret, Shanna H. Swan, Jane Lipscomb, David M. Epstein, Raymond R. Neutra and Laura Fenster, 1992. Spontaneous Abortions and Birth Defects Related to Tap and Bottled Water Use, San Jose, California 1980-1985. *Epidemiology* 3(2): 98-103.

Yang, Chun-Yuh, Bi-Hua Cheng, Shang-Shyue Tsai, Trong-Neng Wu, Meng-Ciao Lin and Kuo-Cherng Lin, 2000. Association Between Chlorination of Drinking Water and Adverse Pregnancy Outcome in Taiwan. *Environmental Health Perspectives* **108**(8).

Chapter 3

A Discussion of Exposure Misclassification

Introduction

For several decades, epidemiologists have investigated a possible association between chemicals in treated drinking water and adverse reproductive effects. However, the ability to determine clear associations has been hampered by exposure misclassification. At several recent workshops[1] on this topic, the question was raised: how good must the exposure assessment need to be, in the absence of biomarkers, to provide an accurate assessment of the strength of the association? This study will address that question from multiple angles, using the methods introduced in this chapter. These are methods for predicting quantitatively the effect of misclassification of categorical and continuous exposure data on the strength of association. Disinfection byproduct (DBP) data original to this study from the water treatment plants and distribution systems of two cities, City A and City B, will be used as well as monitoring data supplied by the water utilities in City A, City B, and City C.

As mentioned in Chapter 1, there are many factors contributing to misclassification of exposure to DBPs in treated drinking water. Predicting the effect of all these factors is beyond the time and budget scope of this study. For this reason, the primary focus of this research is on exposure misclassification as a result of factors affecting DBP concentrations in water. Several DBP species will be investigated.

What is misclassification?

In epidemiology studies, study subjects must be classified in terms of both exposure and outcome: does the subject exhibit the outcome, yes or no; has the subject been exposed and to what extent? Errors in describing the subject in terms of either exposure or outcome can occur due to chance, bias, or confounding. Chance and confounding will not be discussed here. However, one type of bias, called misclassification, is of great interest in this discussion. False positives, where a subject is classified as having an exposure or outcome when in reality she does not, and false negatives, where a subject is classified as not having an exposure or outcome when in reality she does, are both examples of misclassification. We will deal exclusively with exposure misclassification here. The degree of exposure misclassification depends on the accuracy with which exposure measurements are made and assigned to each subject. Epidemiology studies investigating a link between exposure to agents in treated drinking water and adverse reproductive effects have been plagued by exposure misclassification. Elements contributing to exposure misclassification in these studies were discussed in Chapter 1, and include factors affecting chemical concentrations in treated drinking water, exposure pathways, and identification of the causal agents. This chapter will further investigate exposure misclassification and its effects on the measure of association in epidemiology studies.

---

[1] An International Workshop of Exposure Assessment for Disinfection By-Products in Epidemiologic Studies, Ottawa, Ontario, May 7-10, 2000; Safety of Water Disinfection: Balancing Chemical and Microbial risks, Miami, Florida, November 1999.

One type of error that must be mentioned here is Berkson type error. Berkson type error applies when group average exposure is used in place of individual values (Armstrong, 1998). As such, Berkson type error is a potential concern in these epidemiology studies. Berkson error can reduce the precision and power of a study. However, it causes little or no bias, and therefore little or no attenuation of the OR. In this thesis, the focus is exclusively on quantifying OR attenuation as a result of exposure misclassification, a type of bias. In addition, the statistical methods introduced in this chapter and used in the balance of the thesis assume that the only error present in the exposure assessment is exposure misclassification. Because of the fact that we are interested exclusively in the effect of exposure misclassification on OR attenuation and Berkson type error affects the power and precision of a study, not the OR attenuation, Berkson type error will not be considered further in this thesis. However, Berkson type error is a potential problem in these epidemiology studies and should be considered in future work.

Exposure misclassification is described as differential or non-differential based on whether the misclassification is the same or different for those with and without the outcome. Non-differential or random misclassification of exposure occurs when misclassification of the exposure measure is not related to the outcome or disease (Hennekens and Buring, 1987). For example, in a case-control study looking at smoking and risk of lung cancer, 90% of the cases (people with the disease) who were smokers reported their smoking history accurately and 90% of the controls (people without the disease) who were smokers also reported their smoking history accurately. The same proportion of the groups being compared (people with and without lung cancer) reported their smoking history accurately, so the misclassification is said to be independent of the disease state, or non-differential. However, if 90% of cases who smoked reported their smoking history accurately compared to 60% of the controls, then the exposure misclassification cannot be said to be independent of the disease and the exposure would be considered differentially or non-randomly misclassified. The same example can be used to illustrate differential and non-differential misclassification for those study subjects who do not smoke (i.e. are not exposed).

Exposure data can be either continuous or categorical. If subjects are grouped in categories depicting certain exposure ranges, the exposure data are said to be categorical. Categorical exposure can be qualitative or quantitative. For example, a qualitative dichotomous (two-way) exposure assessment could define categories to represent exposure to chlorinated *vs.* non-chlorinated water, or "high" *vs.* "low" exposure to certain chemicals. However, if the concentration ranges for "low" and "high" exposure were defined, for example ≤99 μg/L and ≥100 μg/L, the exposure assessment would be quantitative. Data can be grouped into more than two categories or numerically defined ranges, for example, "high", "medium" and "low", or 0-24 µg/L, 25-49 µg/L, and 50+ µg/L. Continuous data are not grouped into categories, but cover the entire range of possible exposures. Continuous data are always quantitative, and subjects are assigned a discrete exposure value rather than an exposure range or category.

Epidemiological Measures of Strength of Association

Epidemiological statistics are often done on categorical data using a tabulated summary of the exposure and outcome data. These tables are used to calculate a measure of strength of association, either a *relative risk* (RR) or an *odds ratio* (OR). Odds ratios are the main measure used in the epidemiology studies of interest here. Relative risks will not be discussed further. The odds ratio measures the likelihood of exposure between those who have the disease and those who do not (Hennekens and Buring, 1987). The greater the odds ratio, the greater the likelihood of an association between the exposure and the disease. An odds ratio of 1.0 indicates that there is no observed association between the exposure and the disease. The odds ratio is most simply explained by using a two-by-two table shown in Table 3-1.

Table 3-1. Example two-by-two table for the calculation of odds ratio

	Outcome	
Exposure	**Yes**	**No**
Yes	45 (a)	75 (b)
No	5 (c)	75 (d)

The odds ratio is the odds of exposure among those who have the outcome divided by the odds of exposure among those who do not have the outcome (Hennekens and Buring, 1987):

$$OR = (a/c) / (b/d) = ad / bc$$

$$OR = (45*75) / (75*5) = 9$$

Odds ratios can also be calculated for studies employing continuous data. However, the methods for calculating the strength of association using continuous data are quite different from those using categorical data, since continuous data can not be summarized in odds ratio tables. Odds ratios for continuous data are calculated using a logistic regression model. Logistic regression may be used to predict the odds ratios of categorical data as well. An explanation of logistic regression methods for calculating odds ratios can be found in Appendix 1.

Indices of accuracy of measurement

Exposure misclassification is an important element to consider when calculating the odds ratio of an epidemiology study. It is necessary to have an idea of the accuracy of the exposure measures in order to determine the extent of misclassification. There are several indices describing the accuracy of exposure measurements, which are different for categorical and continuous data. These indices of accuracy will aid in determining the effect of misclassification on the results of epidemiology studies.

Categorical data

The accuracy of dichotomous, categorical data is determined by two measures: *sensitivity* and *specificity*. Sensitivity is a measure of the proportion of study subjects who are truly exposed and are classified as exposed. Specificity is a measure of the proportion of study subjects who are truly not exposed and are classified as not exposed. Further explanation can be found in Table 3-2 and the calculations following.

Table 3-2. Example of a table to determine the sensitivity and specificity of a study

	True exposure		
Imperfect exposure	**Yes**	**No**	**Total**
Yes	100 (a)	5 (b)	105
No	10 (c)	85 (d)	95
Total	110	90	

Sensitivity is the proportion of those who are truly exposed who are measured as exposed (Hennekens and Buring, 1987):

$$\text{Sensitivity} = a / (a+c) = 100/110 = 0.91, \text{ or } 91\%$$

Specificity is the proportion of those who are truly not exposed who are measured as not exposed (Hennekens and Buring, 1987):

$$\text{Specificity} = d / (b+d) = 85/90 = 0.94, \text{ or } 94\%$$

It is difficult to calculate sensitivity and specificity for actual epidemiologic studies. Because the imperfect measure is employed in these studies, knowledge of the true exposure measures is seldom known. In most cases, true exposure for the study population can only be determined at great expense, if at all. For some studies, it is possible to measure the true exposure for a small subset of the population in order to give an estimate of the sensitivity and specificity of the exposure data for the whole study.

Continuous data

The accuracy of continuous data can be described by two measures. The first measure is *bias, b*, or the average measurement error in the population (Kelsey et al., 1986). Bias is also referred to as the systematic error in the exposure measurement (Armstrong et al., 1992). Bias can be measured as the difference between the population mean of the imperfect exposure measurements (μ_X) and the population mean of the true exposure measurements (μ_T) (Kelsey et al., 1986).

$$b = \mu_X - \mu_T$$

Bias is sometimes expressed in relation to the total variability and is termed the *standardized bias*. The standardized bias is calculated as follows:

$$\text{Standardized bias} = (\mu_X - \mu_T) / \sigma_T$$

Where σ_T is the standard deviation of the true measurements.

The second measure of the accuracy of the data is the correlation of the imperfect measurements (X) with the true measurements (T). This correlation, depicted by ρ_{TX} is also called the *validity coefficient*. The validity coefficient as used in this context is similar to the statistical Pearson product moment correlation coefficient (r). A related measure is the proportion of the variance of X explained by T, depicted as ρ_{TX}^2. This measure is similar to the statistical coefficient of determination (r^2). The validity coefficient or correlation coefficient, ρ_{TX}, ranges between -1 and 1, with a value of 1 indicating that X is a precise measure of T (Armstrong et al., 1992).

The effect of measurement error on the distribution of the imperfect exposure data relative to the true exposure data can be seen in Figure 3-1. The assumption is made here that the data sets are normally distributed. It can be seen from this figure that the bias shifts the mean of the measured distribution away from the mean of the true distribution. In addition, the decrease in the precision of the data increases the spread of the measured distribution relative to the true distribution.

As with the sensitivity and specificity, the true exposure measures must be known to calculate the bias and correlation.

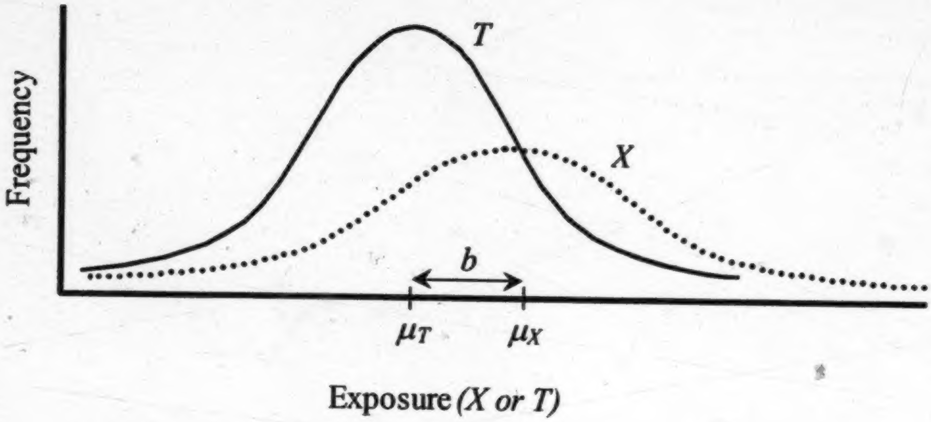


Figure 3-1. Effect of measurement error on the distribution of a normally distributed exposure, after Armstrong et al., 1992.
T = the true exposure; X = the measured exposure; μ_T=the population mean of T; μ_X=the population mean of X; b = the systematic error, or bias, that can occur for all measured values (X).

Effect of misclassification on odds ratios

Categorical Data

Misclassification is an important consideration in evaluating the strength and accuracy of an epidemiologic study. The effect of exposure misclassification on the rate ratios for categorical data can be easily demonstrated using 2x2 tables similar to those introduced previously for calculating odds ratios and employing the concepts of sensitivity and specificity. Table 3-3 shows dichotomous exposure and outcome data and odds ratios (OR) for a hypothetical case-control study.

Table 3-3. Effects of non-differential exposure misclassification on dichotomous exposure variables (Rothman and Greenland, 1998)

Section A: Correct, non-misclassified data

	Exposed	Unexposed
Cases	240	200
Controls	240	600

OR = 3.0

Section B: Sensitivity=0.8, Specificity=1.0

	Exposed	Unexposed
Cases	192	248
Controls	192	648

OR = 2.6

Section C: Sensitivity=0.8, Specificity=0.8

	Exposed	Unexposed
Cases	232	208
Controls	312	528

OR = 1.9

Section D: Sensitivity=0.4, Specificity=0.6

	Exposed	Unexposed
Cases	176	264
Controls	336	504

OR = 1.0

Section E: Sensitivity=0.0, Specificity=0.0

	Exposed	Unexposed
Cases	200	240
Controls	600	240

OR = 0.33

Section A shows the true odds ratio if there is no exposure misclassification. Only non-differential exposure misclassification is considered here and it is assumed that no other errors in subject classification have occurred. Section B of Table 3-3 show the data and calculated OR if all truly unexposed subjects are classified as not exposed, but only 80% of the exposed subjects are correctly classified as exposed. In other words, the sensitivity of the exposure measurement is 80% and the specificity is 100%. Since non-differential classification is assumed, the sensitivity and specificity are the same for both cases and controls in all the examples. This means that of the 240 cases and 240 controls that are truly exposed, only 192 cases and 192 controls (240 x 0.8) are classified as exposed. That means that 48 additional cases and 48 additional controls are incorrectly classified as not exposed, increasing the total number of cases and controls classified as not exposed to 248 and 648, respectively (48 + number of cases or controls truly not exposed). In section B of Table 3-3 a specificity of 80% is seen to lower the OR from 3.0 to 2.6 in this example. In section C, both the specificity of the exposure data and the sensitivity are 80%, decreasing the OR to 1.9. Section D shows exposure data with sensitivity at 40% and specificity at 60%, resulting in an OR of 1.0. It is interesting to note that the non-differential exposure misclassification has completely eliminated any observed positive association between the exposure and the outcome. Section E of the table shows an extreme case where both the specificity and the sensitivity of the exposure data are at 0%. Misclassification this extreme is unlikely to occur in actual studies; however it illustrates the point that extreme misclassification can *reverse* the observed association between an exposure and an outcome. In this case, assuming the outcome is a disease or an unwanted event, the extreme misclassification has resulted in an observed protective effect rather than a significant adverse effect from the exposure.

In a report for Health Canada (Reif et al., 2000) as well as in a presentation at a recent Health Canada workshop[2], an alternate method of predicting the effects of the misclassification of categorical exposure data on the outcome of reproductive epidemiology studies was presented. This work calculates "true" odds ratios from the measured odds ratios and the misclassified data. The analysis was divided into dichotomous exposure categories and multi-category exposure. Only the results related to dichotomous exposure will be presented here. The method begins with a series of equations defining "true" exposure in terms of the measured exposure.

---

[2] An International Workshop of Exposure Assessment for Disinfection By-Products in Epidemiologic Studies, Ottawa, Ontario, May 7-10, 2000

Two 2-by-2 tables are presented:

Classified or measured exposure

Exposed	Cases	Controls
Yes	a	b
No	c	d

"True" exposure

Exposed	Cases	Controls
Yes	A	B
No	C	D

With:
a = cases who are classified as exposed
b = controls who are classified as exposed
c = cases who are classified as not exposed
d = controls who are classified as not exposed
A = cases who are truly exposed
B = controls who are truly exposed
C = cases who are truly not exposed
D = controls who are truly not exposed

Relating these values to sensitivity and specificity, the following equations were developed:
$a = Se_1A + (1-Sp_1)C$
$b = Se_0B + (1-Sp_1)D$
$c = (1-Se_1)A + Sp_1C$
$d = (1-Se_0)B + Sp_0D$

Where:
Se_1 = the sensitivity among cases
Se_0 = the specificity among controls
Sp_1 = the specificity among cases
Sp_0 = the specificity among controls

The total number of cases and controls is not affected by exposure categorization; therefore, a+c = A+C, and b+d = B+D.

The "true" exposure classifications for a ,b , c, and d can be calculated as follows:
$A = [Sp_1a - (1-Sp_1)c]/[Se_1Sp_1 - (1-Se_1)(1-Sp_1)]$
$B = [Sp_0b - (1-Sp_0)d]/[Se_0Sp_0 - (1-Se_0)(1-Sp_0)]$
$C = a+c-A$
$D = b+d-B$

Using these values for A, B, C, and D, the "true" OR can be calculated. Comparing this calculated "true" OR with the observed OR will give an estimate of the effect of exposure misclassification on the OR of the study. In order to calculate the sensitivity and specificity of the exposure data in a study, the true exposure data must be known for at least a subpopulation of the study subjects. However, if the actual sensitivity and specificity cannot be calculated for the study, hypothetical sensitivities and specificities provide a useful thought experiment.

Reif et al. (2000) applied these equations to real data from two epidemiology studies on neural tube defects (Bove et al., 1992; Klotz and Pyrch, 1999). The results are shown in Tables 3-4 and 3-5.

Table 3-4. Calculated "true" ORs for the Klotz and Pyrch (1999) study at different levels of sensitivity and specificity[a]

		Specificity				
		1.00	0.95	0.90	0.85	0.80
Sensitivity	1.00	1.79	1.98	2.35	3.35	14.61
	0.95	1.81	2.00	2.38	3.38	14.78
	0.90	1.83	2.03	2.41	3.43	14.97
	0.85	1.86	2.06	2.45	3.48	15.20
	0.80	1.89	2.10	2.49	3.54	15.48
	0.75	1.93	2.14	2.55	3.62	15.82
	0.70	1.99	2.20	2.62	3.72	16.25
	0.65	2.05	2.28	2.71	3.85	16.82
	0.60	2.15	2.38	2.83	4.03	17.58

[a] a=36, b=52, c=76, d=196; Observed odds ratio: 1.785; 95% CI: 1.08-2.95

Table 3-5. Calculated "true" ORs for the Bove et al. (1992) study at different levels of sensitivity and specificity[b]

		Specificity				
		1.00	0.95	0.90	0.85	0.80
Sensitivity	1.00	1.34	1.41	1.52	1.76	2.50
	0.95	1.35	1.42	1.53	1.77	2.52
	0.90	1.36	1.43	1.54	1.78	2.54
	0.85	1.37	1.44	1.56	1.79	2.56
	0.80	1.39	1.45	1.57	1.81	2.58
	0.75	1.43	1.47	1.59	1.83	2.61
	0.70	1.42	1.49	1.61	1.86	2.65
	0.65	1.45	1.52	1.64	1.89	2.70
	0.60	1.49	1.56	1.68	1.94	2.77

[b] a=17, b=12814, c=39, d=39520; Observed odds ratio: 1.344; 95% CI: 0.76-2.38)

The tables show that even a slight misclassification, represented as changes in specificity and sensitivity, can have a large effect on the OR of a study. Slight misclassification can increase the difference between the "true" OR and the measured OR. In these results, the measured OR is the OR calculated in the published study, with the "true" OR calculated using the equations above, assuming values for the sensitivity and the specificity. An interesting result seen in these tables is that specificity has a greater effect on the measured OR than does sensitivity. It is evident from the tables that misclassification in these two cases alters the OR towards the null value.

The "true" OR values calculated by assuming certain levels of misclassification suggest a significant effect on the OR. The published OR for the Klotz and Pyrch (1999) study had a calculated OR of 1.785 (95% CI 1.08-2.95). This association is statistically significant, but is fairly weak. However, at specificity and sensitivity of 90% the calculated "true" OR is 2.409 (95% CI 1.39-4.17). This is a much stronger association. The Bove et al. (1992) study shows similar results. The calculated OR for this study was 1.344 (95% CI 0.76-2.38), an association that is not statistically significant. With sensitivity and specificity at 90%, the calculated "true" OR is 1.544 (95% CI 0.85-2.82), while a statistically significant association is seen with sensitivity and specificity at 80% with a calculated "true" OR of 2.582 (95% CI 1.29-5.17).

There is an anomaly in the data due to the mathematical relationship of the equations above. The Klotz and Pyrch (1999) study shows very high "true" ORs when the specificity is at 80%; however, the Bove et al. (1992) does not show a similar result. In both studies the ratios of cases classified as unexposed to cases classified as exposed is approximately 2:1. However, the ratios of controls classified as unexposed to controls classified as exposed is approximately 3:1 for the Bove et al. (1992) study and approximately 4:1 for the Klotz and Pyrch (1999) study. Going back to the equations and assuming non-differential misclassification (sensitivity and specificity are the same for cases and controls), the "true" number of exposed controls can be re-written as (Reif et al., 2000):

$$B = Spb + Spd - d$$

Where b is the number of controls classified as exposed and d is the number of controls classified as unexposed. Substituting d as a multiple of b (d=mb) gives:

$$B = (m+1)Spb - mb$$

This equation shows that B approaches zero as m approaches Sp/(1-Sp). As B approaches zero, the calculated "true" OR will increase (from the calculation, OR = AD/BC). For m = Sp/(1-Sp), B will equal zero. In the Klotz and Pyrch example, the ORs become inflated when the Sp = 80%. B will equal zero when m equals four (m = 0.8/0.2). In other words, if there are four times as many controls classified as unexposed (d) as there are controls classified as exposed (b), then the "true" number of exposed controls (B) will be zero, an unlikely occurrence. However, note that in the Klotz and Pyrch study, the ratio of controls classified as unexposed (d) to those classified as exposed (b) is 4:1; the same as the ratio calculated when specificity equals 80%, resulting in m=4. As a result, the calculation of the "true" OR depends in part on the distribution of the study subjects in the exposure categories. Therefore, it can be postulated that with a ratio in an actual study of unexposed classified controls to exposed classified controls of 4:1, the specificity is unlikely to be true and the resulting calculated "true" ORs are artificially inflated due to the mathematics of the method.

Armstrong et al. (1992) also present a method for predicting the effect of misclassification of categorical data on the measure of association of a study. As with the previous methods, sensitivity and specificity are key when dealing with dichotomous exposure.

In a dichotomous exposure, a proportion of subjects from both the truly exposed and truly unexposed categories may be classified as exposed. This is the proportion of the subjects who are *observed* to be exposed (p) and can be expressed as:

$$p = \text{sensitivity}^* P + (1\text{-specificity})^*(1\text{-}P)$$

Where P is the proportion of subjects *truly* exposed.
In addition, two tables can be constructed to determine the true and observed ORs using the proportions of subjects exposed:

True Classification

	Disease	
Exposure	Yes	No
Yes	P_D	P_N
No	$1\text{-}P_D$	$1\text{-}P_N$

$$OR_T = [P_D(1\text{-}P_D)]/[P_N(1\text{-}P_N)]$$

OR_T is the true OR, P_D is the true proportion of subjects exposed in the diseased group, P_N is the true proportion of subjects exposed in the non-diseased group.

Measured Classification (misclassified)

	Disease	
Exposure	Yes	No
Yes	p_D	p_N
No	$1\text{-}p_D$	$1\text{-}p_N$

$$OR_X = [p_D(1\text{-}p_D)]/[p_N(1\text{-}p_N)]$$

OR_X is the measured or observed OR, p_D is the proportion of subjects classified as exposed in the diseased group, p_N is the proportion of subjects classified as exposed in the non-diseased group.

To present the method of determining the effect of exposure misclassification on the OR of a study, the case of differential misclassification must be examined first. In differential exposure misclassification, either the sensitivity of the diseased and non-diseased groups differ, or the specificity of the diseased and non-diseased groups differ, or both. Applying the equation for the proportion of subjects classified as exposed gives:

$$p_D = sens_D P_D + (1\text{-}spec_D)(1\text{-}P_D)$$
$$p_N = sens_N P_N + (1\text{-}spec_N)(1\text{-}P_N)$$
$$OR_X = [p_D(1\text{-}p_N)]/[p_N(1\text{-}p_D)]$$

Where $sens_N$ and $sens_D$ are the sensitivity for the non-diseased and diseased groups, respectively, and $spec_N$ and $spec_D$ are the specificity for the non-diseased and diseased groups, respectively.

For non-differential misclassification, the equations above are valid, but the sensitivity and specificity for the diseased and non-diseased groups are the same ($sens_D = sens_N$, and $spec_D = spec_N$). From these equations, Table 3-6 can be developed to predict the effect of non-differential misclassification on the OR of a study, given the true OR (OR_T) and the sensitivity and specificity of the exposure data. Table 3-7 given here is a much shortened form of Table 3-6 for illustration purposes. Table 3-6 can be found in full in the Appendix for Chapter 3. P_N in both tables is the proportion of exposed subjects who do not have the disease and is determined arbitrarily in the table. P_D, another value required for the development of the tables, is determined as (rearranging the equation for OR_T):

$$P_D = P_N OR_T/[1+P_N(OR_T\text{-}1)]$$

Figures 3-2 to 3-5 are graphical representations of the data in Table 3-7.

Table 3-7 Selected data from Table 3-6 (Appendix): effects of misclassification on categorical data, $P_N = 0.2$

Specificity	Sensitivity	True OR			
		1.20	3.50	7.00	10.0
0.8	0.1	0.98	0.83	0.72	0.67
0.8	0.2	1.00	1.00	1.00	1.00
0.8	0.3	1.02	1.16	1.27	1.32
0.8	0.4	1.03	1.31	1.54	1.65
0.8	0.5	1.05	1.47	1.83	2.01
0.8	0.6	1.06	1.62	2.14	2.43
0.8	0.7	1.07	1.78	2.51	2.94
0.8	0.8	1.09	1.96	2.96	3.60
0.8	0.9	1.10	2.16	3.53	4.53
0.8	0.95	1.11	2.27	3.90	5.17
0.8	0.99	1.11	2.36	4.24	5.81
0.1	0.8	0.97	0.79	0.70	0.66
0.2	0.8	1.00	1.00	1.00	1.00
0.3	0.8	1.02	1.15	1.26	1.31
0.4	0.8	1.03	1.27	1.50	1.63
0.5	0.8	1.04	1.40	1.76	1.96
0.6	0.8	1.05	1.54	2.05	2.36
0.7	0.8	1.07	1.71	2.43	2.88
0.8	0.8	1.09	1.96	2.96	3.60
0.9	0.8	1.12	2.36	3.80	4.75
0.95	0.8	1.15	2.76	4.46	5.66
0.99	0.8	1.18	3.02	5.21	6.68

Table 3-7 continued. Selected data from Table 3-6 (Appendix): effects of misclassification on categorical data, $P_N = 0.2$

Specificity	Sensitivity	True OR			
		1.20	3.50	7.00	10.0
0.99	0.1	1.10	1.90	2.50	2.79
0.99	0.2	1.13	2.17	2.99	3.38
0.99	0.3	1.14	2.33	3.31	3.80
0.99	0.4	1.15	2.46	3.61	4.20
0.99	0.5	1.16	2.59	3.92	4.65
0.99	0.6	1.17	2.72	4.27	5.17
0.99	0.7	1.17	2.86	4.69	5.82
0.99	0.8	1.18	3.02	5.21	6.68
0.99	0.9	1.19	3.20	5.88	7.87
0.99	0.95	1.19	3.30	6.29	8.66
0.99	0.99	1.19	3.38	6.67	9.44
0.1	0.99	1.04	1.45	2.00	2.41
0.2	0.99	1.04	1.54	2.25	2.81
0.3	0.99	1.05	1.62	2.45	3.12
0.4	0.99	1.06	1.70	2.65	3.44
0.5	0.99	1.07	1.81	2.90	3.81
0.6	0.99	1.08	1.94	3.21	4.28
0.7	0.99	1.09	2.11	3.63	4.91
0.8	0.99	1.11	2.36	4.24	5.81
0.9	0.99	1.14	2.76	5.19	7.23
0.95	0.99	1.17	3.06	5.90	8.29
0.99	0.99	1.19	3.38	6.67	9.44

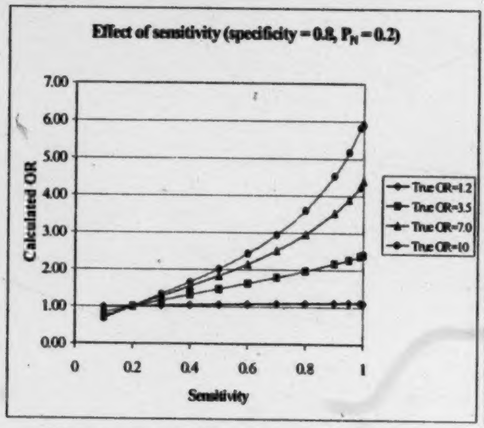


Figure 3-2. Effect of sensitivity

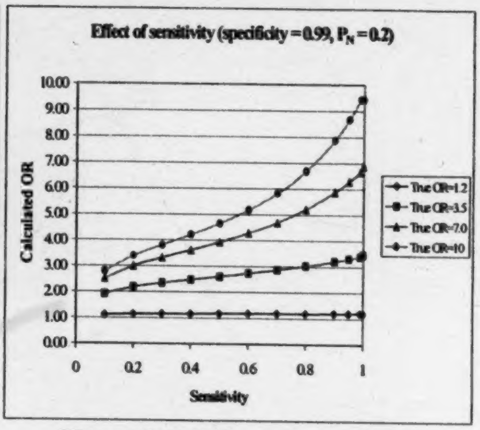


Figure 3-3. Effect of sensitivity

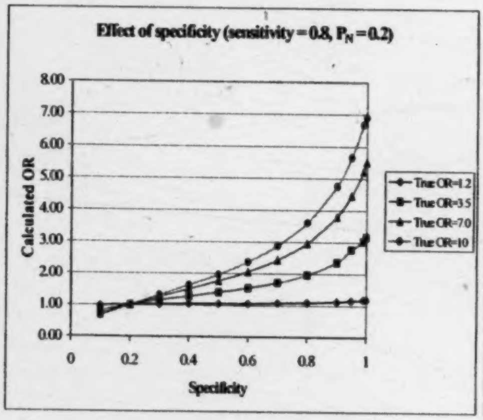


Figure 3-4. Effect of specificity

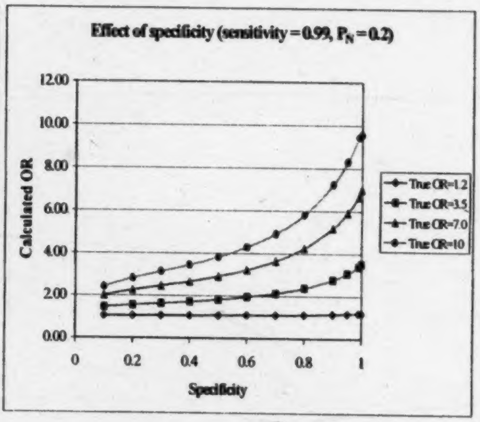


Figure 3-5. Effect of specificity

A reduction in the accuracy of the measured exposure data as expressed by the sensitivity and specificity, attenuates the observed OR towards the null value and in some cases even beyond the null value. Changes in specificity when sensitivity is kept constant seem to result in more steeply sloped lines than changes in sensitivity when specificity is constant. This supports previous findings that specificity has a greater effect on the OR than does sensitivity. In addition, as found in Reif et al. (2000), the observed OR (OR_X) depends not only on the sensitivity and the specificity of the exposure data, but also on the distribution of study subjects (P_N). These tables can be used to predict the effect of misclassification on the OR of studies using categorical exposure data.

There is an anomaly in Tables 3-6 and 3-7, illustrated in Table 3-7 when specificity=0.8 and sensitivity=0.2 or *vice versa*. At these points in Table 3-7, the calculated OR are all 1.00. This same result can be seen in Table 3-6 and occurs wherever (sensitivity = 1-specificity). The reason for this can be seen by substituting (sensitivity=1-specificity) into the equations used to calculate the OR.

If sensitivity=1-specificity, then
$p_D = \text{sens } P_D + ((1\text{-spec})(1-P_D) = \text{sens}P_D + \text{sens }(1-P_D) = \text{sens }(P_D + 1 - P_D) = \text{sens}$
and
$p_N = \text{sens } P_N + (1\text{-spec})(1-P_N) = \text{sens}P_N + \text{sens}(1-P_N) = \text{sens }(P_N + 1 - P_N) = \text{sens}$
So, $p_N = \text{sens} = p_D$

Then, substituting $p_N = p_D$ gives
$OR_X = [p_D(1-p_N)]/[p_N(1-p_D)] = [p_D(1-p_D)]/[p_D(1-p_D)]$ {or $[p_N(1-p_N)]/[p_N(1-p_N)]$} = 1

Due to this mathematical anomaly, it is not possible to predict the OR of a study if (sensitivity = 1-specificity).

Bias away from the null
The examples of the effect of non-differential misclassification described above all resulted in a bias towards the null value and possibly beyond the null value to reverse the observed effect. Attenuation towards the null is the expected effect of non-differential exposure misclassification. However, if categorical data are divided into more than two categories, it is possible that exposure misclassification will result in an exaggeration of the effect and an observed OR biased *away* from the null value. An example where this could occur is presented in Table 3-8, using data from another hypothetical epidemiology study.

Table 3-8: Effects of non-differential exposure misclassification on polychotomous exposure variables (Rothman and Greenland, 1998)

Correct, non-misclassified data

	Unexposed	Low Exposure	High Exposure
Cases	100	200	600
Controls	100	100	100
OR	1	2	6

40% of high exposure misclassified as low exposure

	Unexposed	Low Exposure	High Exposure
Cases	100	440	360
Controls	100	140	60
OR	1	3.1	6

In this case, non-differential exposure misclassification has occurred in that 40% of the subjects in the high exposure category have been misclassified into the low exposure category. None of the unexposed subjects or the truly low-exposed subjects have been misclassified. The resulting OR for the low exposure category is 3.1, or away from the null value compared to the true OR of 2.0 for that category. The OR for the high exposure category did not change.

Another example of non-differential exposure misclassification resulting in a stronger association is presented in Table 3-9.

Table 3-9. Another example of the possible effects of non-differential misclassification of polychotomous exposure variables (Wacholder et al., 1991)

Correct, non-misclassified data

	None	Low	High	Low or high
Cases	500	200	300	500
Controls	700	280	70	350
Odds Ratios	1.00	1.00	6.00	2.00

Misclassified data

	None	Low	High	Low or high
Cases	540	160	300	460
Controls	756	224	70	294
Odds Ratios	1.00	1.00	6.00	2.19

In this example, 20% of the low exposure subjects have been misclassified as not exposed. The non-differential exposure misclassification only affects the OR when the low and high exposure categories are collapsed into a single "any exposure" category. However, in this new single category, the observed effect is stronger than the true effect calculated with no misclassification.

Continuous Data

When the misclassification of interest in continuous data is non-differential, there is equal bias between the diseased and the non-diseased groups. The difference in the mean of the imperfect exposure measurement values between the diseased (D) and non-diseased (N) subjects (μ_{XD}-μ_{XN}) equals the difference in the means of the true exposure measurement values between the diseased and non-diseased subjects (μ_{TD}-μ_{TN}). This is shown graphically in Figure 3-6, where $\mu_{XN} = \mu_{TN}$ and $\mu_{XD} = \mu_{TD}$; therefore, $\mu_{XD}-\mu_{XN} = \mu_{TD}-\mu_{TN}$.

Also shown in Figure 3-6 is the widening of the distribution of X relative to T as a result of the lack of precision in X. Accordingly, there is more overlap and less distinction between X_D and X_N than there is between T_D and T_N. This results in an OR curve flattened towards the null value as shown in Figure 3-7. The null value is represented by the horizontal line of OR = 1 for all exposures (Figure 3-7).

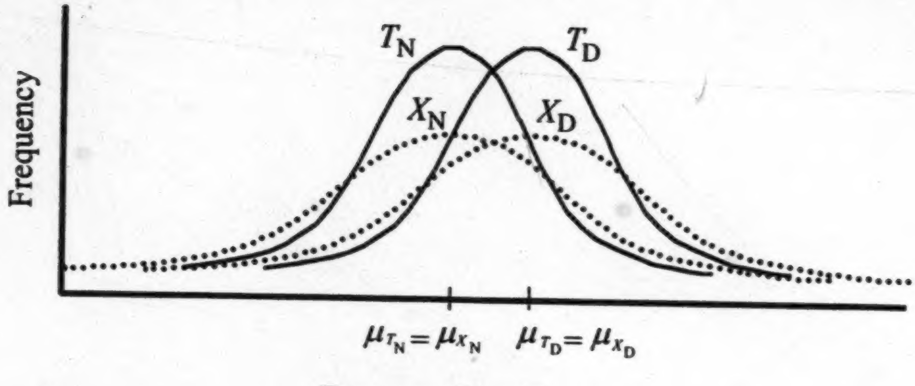


Figure 3-6. Effect of non-differential measurement error on the exposure distributions of the non-diseased and diseased groups, after Armstrong et al., 1992
Where: T_N and T_D are the true exposures in the non-diseased and diseased groups, respectively; X_N and X_D are the measured exposures in the non-diseased and diseased groups, respectively

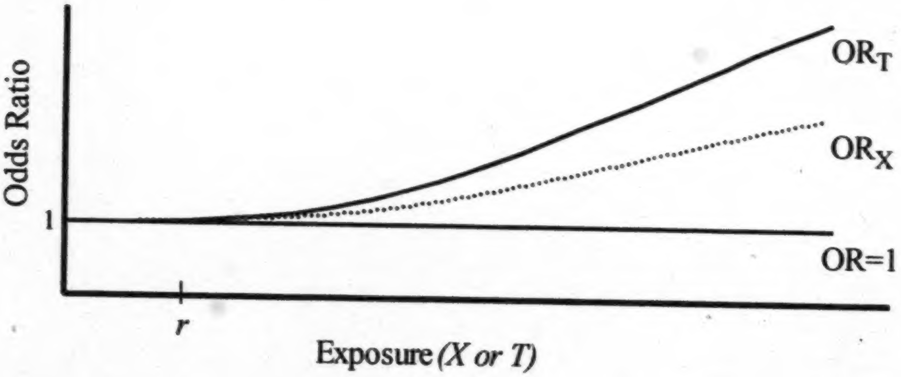


Figure 3-7. Effect of non-differential measurement error on the odds ratio curve, after Armstrong et al., 1992
Where: OR_T=the true OR, OR_X=the measured, attenuated OR, r=the reference level

Previously, a measure of precision of X was introduced as the correlation of T with X. This measure was called the validity coefficient. The model for the attenuation equation, introduced below, suggests that the square of the validity coefficient is related to the variances of X, T, and E (the subject error) as follows:

$$\rho^2_{TX} = 1 - \sigma^2_E/\sigma^2_X = \sigma^2_T/\sigma^2_X$$

Mathematical proofs for these relationships can be found in the Appendix.
With this relationship, the mathematical basis for Figure 3-6 becomes clear. The equation can be rewritten as:

$$\sigma^2_X = \sigma^2_T/\rho^2_{TX}$$

This expression shows that the variance of X is expected to be greater than the variance of T, since by definition $0 \leq \rho^2_{TX} \leq 1$.

To describe the effect of misclassification of the continuous data numerically, it is necessary to introduce the *attenuation equation*:

$$\beta_X = \rho^2_{XT}\beta_T$$

Where β_X and β_T are the logistic regression coefficients (see Appendix: Logistic Regression Primer) calculated from the imperfect data and the true data, respectively. The attenuation equation indicates that βx tends towards the null value as the accuracy in the measured data (ρ_{XT}) decreases.

The attenuation equation is based on a model of measurement error in a population where the observed exposure (X) for an individual, *i*, is equal to the true exposure (T) added to the systematic error or bias (b), and the subject error (E):

$$X_i = T_i + b + E_I$$

The subject error (E) is the error in exposure measurement that varies from subject to subject. It is assumed in this model that $\mu_E = 0$ and $\rho_{TE} = 0$. The second term states that there is no correlation between the true exposure and the subject error. An example of $\rho_{TE} \neq 0$ would occur if the subject error increased as the true exposure measure increased. X, T, and E are assumed to be distributions with means (μ_X, μ_T, μ_E) and variances (σ^2_X, σ^2_T, σ^2_E).

The attenuation model and all the equations presented here assume that the sole source of error is exposure misclassification. Errors in the measurement of disease, confounding, sampling errors, and selection bias are assumed not to be present.

The measurement error model can be presented for both the non-diseased (N) and diseased groups (D) as follows, where X is calculated for every subject, i:

$$X_{iN} = T_{iN} + b_N + E_{iN}$$
$$X_{iD} = T_{iD} + b_D + E_{iD}$$

To develop the concept of the attenuation equation, it is necessary to first examine the case of differential misclassification. One of two conditions must be satisfied for differential misclassification to occur: either $b_N \neq b_D$, or the precision of X_N differs from that of X_D ($\sigma^2_{EN} \neq \sigma^2_{ED}$).

There are several assumptions made in order to simplify the presentation of the attenuation equation and the model it is based on:

1) X_N and X_D follow the model presented above and have $\rho_{TE} = 0$
2) T_N and T_D are normally distributed with the same variance ($\sigma^2_{TD} = \sigma^2_{TN} = \sigma^2_T$) and have means of μ_{TN} and μ_{TD}, respectively
3) E_N and E_D are normally distributed with $\mu_E = 0$ and the same variance ($\sigma^2_{EN} = \sigma^2_{ED} = \sigma^2_E$). This assumption establishes that the differential misclassification is a result of differential bias only, and not differential precision.

From the model and assumptions presented above, a logistic regression model can be developed for the probability of disease as a function of the true exposure, T (Wu et al., 1986):

$$\log\left[\Pr(d)/(1-\Pr(d))\right] = \alpha_T + \beta_T T$$

Where Pr(d) is the probability of disease, and β_T is the logistic regression coefficient,
$\beta_T = (\mu_{TD} - \mu_{TN})/\sigma^2_T$.

A similar equation can be written for the measured variable, X:

$$\log\left[\Pr(d)/(1-\Pr(d))\right] = \alpha_X + \beta_X X$$

Where $\beta_X = [(\mu_{TD} - \mu_{TN}) + (b_D - b_N)]/(\sigma^2_T + \sigma^2_E)$.

β_X can then be re-written in terms of β_T as follows (the algebraic proof of this relationship can be found in the appendix):

$$\beta_X = [1 + (b_D - b_N)/(\mu_{TD} - \mu_{TN})]\, \beta_T \rho^2_{TX}$$

This equation gives the relationship between the logistic regression coefficient for the observed exposure data in terms of the logistic regression coefficient for the true exposure data, assuming differential misclassification. However, if non-differential misclassification is considered, the bias between the diseased and non-diseased subjects is equal ($b_D = b_N$), and the equation becomes:

$$\beta_X = \rho^2_{TX}\beta_T$$

This equation, then, is the attenuation equation, which indicates that the logistic regression coefficient for the observed exposure data is attenuated towards the null value of 0 (because $0 \leq \rho^2_{TX} \leq 1$) as the correlation between the true and observed data decreases.

The attenuation can be interpreted in terms of the OR in two ways. The first is in terms of a fixed-unit (k) increase in both the measured (X) and true (T) exposure data (see Appendix information on logistic regression). The true and measured ORs (OR_T and OR_X, respectively) can be described by the following equations:

$$OR_T = \exp(\beta_T k)$$
$$OR_X = \exp(\beta_X k)$$

Then substituting the attenuation equation gives OR_X in terms of OR_T:

$$OR_X = OR_T \exp(\rho^2_{TX})$$

This equation states that the measured OR for any k-units of difference in X equals the true OR for any k-units of difference in T to the exponent ρ^2_{TX}. Since $0 \leq \rho_{TX} \leq 1$, $OR_X \leq OR_T$.

The second way of interpreting the attenuation equation is by comparing the measured and true ORs based on a difference of s-standard deviations of X and T rather than by a k-unit difference in X and T. This method is often used to compensate for the increased spread of the measured exposure data relative to the true exposure data. The true and measured ORs can be described by the following equations:

$$OR_T = \exp(\beta_T s \sigma_T)$$
$$OR_X = \exp(\beta_X s \sigma_X)$$

Where σ_T and σ_X are the standard deviations in T and X, respectively, and s is the number of standard deviations.

OR_X can then be interpreted in terms of OR_T as follows:

$$OR_X = OR_T \exp(\rho_{TX})$$

This equation states that the measured OR for s-standard deviations in X equals the true OR for s-standard deviations in T to the exponent ρ_{TX}. Again, since $-1 \leq \rho_{TX} \leq 1$, $OR_X \leq OR_T$. The attenuation in the measured OR is less when the OR is interpreted in terms of the standard deviation rather than a fixed unit change in the exposure data.

These equations can be used to determine the result of non-differential exposure measurement errors on the strength of association measurement (OR) in epidemiologic studies using continuous data. With knowledge of the validity coefficient (ρ_{TX}) between the measured and true exposure data sets and the OR calculated for the true exposure data, it is possible to predict the OR for the misclassified exposure. The equation for a k-unit change will not be used further in this work. The equations for standard deviation will be employed for all analyses. The standard deviation equation allows for a comparison between data sets with very different ranges. The standard deviation allows for an equivalent measure of change between the data sets to be compared here whereas the unit change approach would be affected by the range of data in each case.

Table 3-10 was developed using the equations presented above. Table 3-10 can be found in the Appendix for Chapter 3. Selected data from Table 3-10 are shown in Table 3-11 for illustrative purposes. Figures 3-8 and 3-9 are graphical representations of the data in Table 3-11.

Table 3-11. Selected data from Table 3-10 (Appendix): effects of misclassification on continuous data

ρ_{XT}	ρ^2_{XT}	OR_Xa	OR_Xb	OR_Xa	OR_Xb	OR_Xa	OR_Xb	OR_Xa	OR_Xb
		1.20		3.50		7.00		10.00	
0.100	**0.010**	1.02	1.00	1.13	1.01	1.21	1.02	1.26	1.02
0.200	**0.040**	1.04	1.01	1.28	1.05	1.48	1.08	1.58	1.10
0.300	**0.090**	1.06	1.02	1.46	1.12	1.79	1.19	2.00	1.3
0.400	**0.160**	1.08	1.03	1.65	1.22	2.18	1.37	2.51	1.45
0.500	**0.250**	1.10	1.05	1.87	1.39	2.65	1.63	3.16	1.78
0.600	**0.360**	1.12	1.07	2.12	1.59	3.21	2.01	3.98	2.29
0.700	**0.490**	1.14	1.09	2.40	1.85	3.90	2.59	5.01	3.09
0.800	**0.640**	1.16	1.12	2.72	2.23	4.74	3.47	6.31	4.37
0.900	**0.810**	1.18	1.16	3.09	2.76	5.76	4.84	7.94	6.46
0.950	**0.903**	1.19	1.18	3.29	3.10	6.35	5.79	8.91	7.99
0.980	**0.960**	1.20	1.19	3.41	3.33	6.73	6.48	9.55	9.13
0.990	**0.980**	1.20	1.20	3.46	3.41	6.87	6.73	9.77	9.55

$OR_Xa = OR_T\exp(\rho_{XT})$ (for s standard deviation change in the exposure data)

$OR_Xb = OR_T\exp(\rho^2_{XT})$ (for any k-unit change in the exposure data)

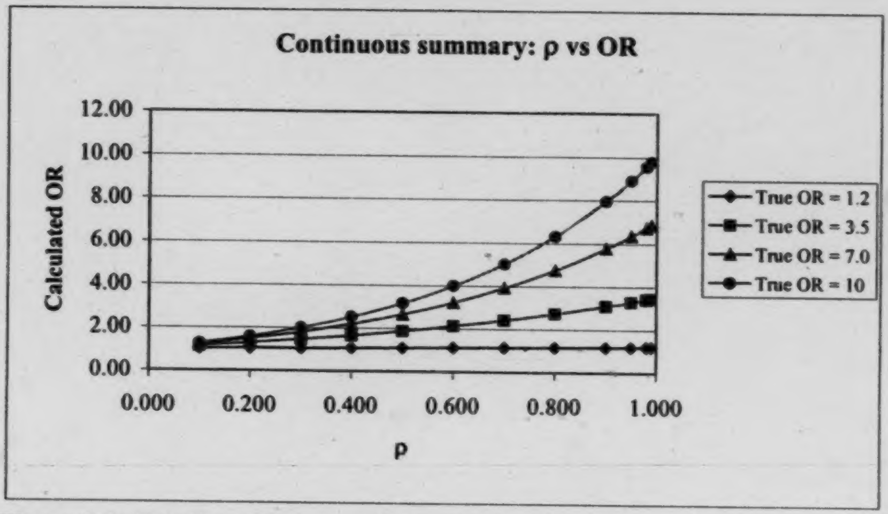


Figure 3-8: Relationship between ρ and OR attenuation

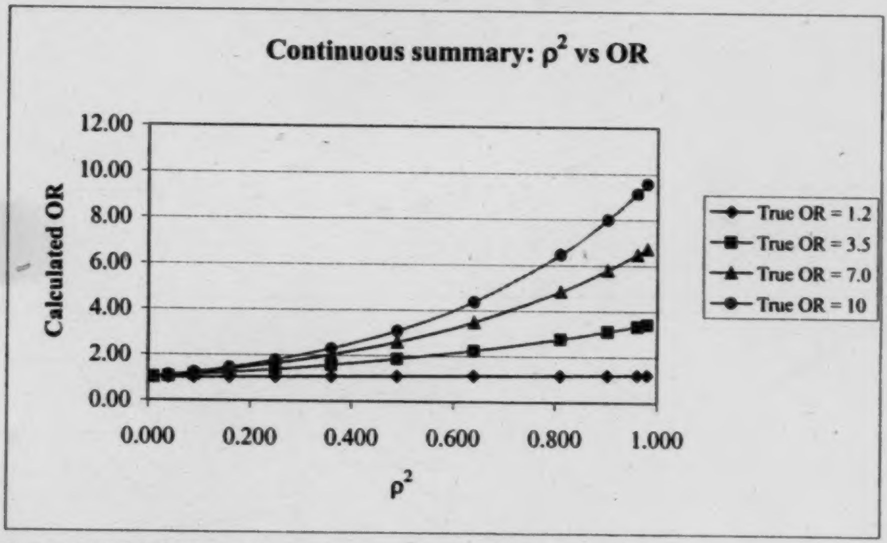


Figure 3-9. Relationship between ρ^2 and OR attenuation

Tables 3-10 and 3-11 and Figures 3-8 and 3-9 show the relationship between the correlation of the true and measured data and the resulting OR for a study using the measured data. In contrast to the method of predicting the OR using categorical data, the calculations using continuous data will not result in an OR less than the null value (OR = 1.0), unless $\rho<0$. This is true even when there is no correlation between the true and measured data ($\rho=0$). Also, very small differences between the measured and true data sets measured as r-values near, but not at, one can have a substantial effect on the resulting OR. This effect is greater the larger the true OR. With a knowledge of both the true and measured exposure data, this method can predict the effect of misclassification on the OR of a study.

Table 3-10 and the attenuation equation hold for a specific error measurement model under certain assumptions. The model given at the beginning of the discussion relates the observed exposure data to the true exposure data with the addition of simple error terms. It is assumed that the magnitude of the measurement error is not related to the magnitude of the true measurement ($\rho_{TE} = 0$). Additionally, it is assumed that the only error leading to exposure misclassification is error in the exposure measurement. Other errors in the identification of disease outcome, selection bias, or confounding are not considered and are assumed to be absent. X, T and E are assumed to be normally distributed. The assumption is also made that the true exposures are known to the extent that a validity coefficient can be calculated. Under this simple model, the effects of exposure misclassification from errors in the measurement of exposure on the results of a study can be estimated.

However, it is possible that the simple model may be violated. A violation of the model may occur if X is a function of T, which could happen if X is a measure of T using a different scale of measurement. In this case the model would be:

$$X_i = cT_i + b + E_i$$

The equation that would still hold under this new model is:

$$OR_X = OR_T\exp(\rho_{TX}).$$

The equation $OR_X = OR_T\exp(\rho^2_{TX})$ would not hold since the derivation of this equation would be affected by the linear transformation of X. This limits the effectiveness of Table 3-10 in terms of the interpretation of an estimation of the effect of misclassification on the OR of an epidemiology study. However, the results of the part of Table 3-9 obtained by using the equation $OR_X = OR_T\exp(\rho_{TX})$ may be used, if all other conditions are satisfied.

Another group (de Klerk et al., 1989), suggests that the true OR of a study can be estimated using the observed OR and the validity coefficient between the measured and true exposure measures, in essence, reversing the equations presented above. While this is a tempting proposition, any estimates of the true OR made in this way must be taken with a grain of salt, since the original equations are based on a simplified model with several assumptions.

Bias away from the null

Random or non-differential misclassification of either categorical or continuous exposure is commonly expected to yield an odds ratio that is biased towards the null value, giving an underestimate of the true risk. However, as was seen in the case of categorical exposure, this is not always the case. Continuous data are often categorized in the statistical treatment of epidemiology studies, due in part to the inherent error in exposure measurement. Exposure assessment resulting in specific numerical measurements of exposure for individual people may be subject to considerable error, as discussed in Chapter 1. For this reason subjects may be grouped into exposure categories based on their continuous exposure measures. However, if non-differential misclassification occurred in assigning continuous exposure values to study subjects, then collapsing the continuous data into categories could introduce an element of differential misclassification into the study (Flegal et al., 1991; Wacholder et al., 1991).

Another group (Delpizzo and Borghesi, 1995) supports the work done by Flegal on collapsing exposure categories. They argue that although non-differential misclassification may occur in exposure assessment, the differential misclassification resulting from the categorization of continuous data provides for the possibility that the measured rate ratio may be biased away from the null value with respect to the true rate ratio. The authors of another study (Dosemeci et al., 1990) argue that non-differential misclassification can bias the measured rate ratio away from the null value in polychotomous variables, but not dichotomous. Essentially, bias away from the null usually only occurs when differential misclassification is introduced into the study by collapsing several exposure categories into fewer categories.

Development of Test Model

A table similar to Table 3-10 is presented in Kelsey et al. (1986) without reference to the equations introduced in Armstrong et al. (1992). A model was developed by the author and A. Senthilselvan to reproduce and expand this table without the benefit of these equations. The model was developed to test whether the results of Table 3-10 could be duplicated without reference to the equations. This model and the resulting table are seen to support the attenuation equation and subsequent equations leading to the formation of Table 3-11. Normally distributed data sets for exposure and outcome were generated in SPSS (Version 10.0.5 for Windows, statistical software by SPSS Inc.). Exposure data sets correlated to the original exposure data set, referred to as the "misclassified" exposure, were also generated in SPSS. The correlation coefficients were calculated via linear regression of the two exposure data sets. ORs were calculated with logistic regression in SPSS using these "misclassified" exposure data. A table similar to Table 3-10 was then constructed with the "true" ORs from the original exposure data, and the ORs calculated from the "misclassified" data, as well as the correlation coefficient between the "true" and "misclassified" exposure. This model has the advantage of being based on very few assumptions. The main assumptions are that the exposure data sets are normally distributed, and that the error between the "true" and "misclassified" exposure data is random. Using the true OR and the correlation coefficient in the OR equations from Armstrong et al. (1992) gives the observed ORs calculated in the model.

Applying the methods for predicting the effect of exposure misclassification on the strength of association of epidemiology studies

At the beginning of this chapter, the question was raised: how good does exposure assessment need to be, in the absence of biomarkers, in order to provide an accurate assessment of the strength of an association between agents in chlorinated drinking water and adverse reproductive effects? The tools introduced in this chapter will allow us to address this question with reasonable assumptions, utilizing data on water concentrations of disinfection by-products (DBPs). The DBP concentration data is continuous, but it will be treated both as continuous and as categorical (dichotomous). The data available for use include original data collected from the distribution system and water treatment plants in City A and City B as part of this study, as well as historical monitoring data provided by the water utilities in City A, City B, and City C. The experimental details are given in Chapter 4.

Daily data: City A

Samples were collected from two water treatment plants (Water treatment plant #1 and Water treatment plant #2) as well as 4 distribution locations (households) in City A. The distribution locations were chosen so as to provide different mixes of water from the two water treatment plants as well as different residence times, estimated in part by distance from the water treatment plants within the distribution system. Samples of cold tap water were collected daily over a one-month period (29 days) and were analysed for TCM, DCAN, DCAA, and TCAA.

Daily data: City B

Samples were collected from one water treatment plant and 4 distribution system locations (households) located at different residence times from the water treatment plant, as estimated by distance from the plant, and pump flow information in the distribution system during the sampling period. Samples of cold tap water were collected daily over a one-month period (29 days) and analysed for TCM, BDCM, DCAA, and TCAA.

Monitoring Data

The water utilities of City A, City B and City C generously shared their monitoring data for use in this project. These data are further described in Chapter 4

Categorical treatment of data

In order to treat the continuous data as dichotomous, a cut-point must be chosen, below which subjects are assumed to be "not exposed" and above which they are assumed to be "exposed". If two or more different measures of the same exposure are treated this way and one is assumed to be the "true" exposure while the others are assumed to be the "observed" exposures, sensitivity and specificity can be calculated for that exposure, measured several different ways. An example of this is shown in the following exercise using published data from a study that compared TCAA concentrations in drinking water with calculated ingestion dose and measured TCAA excretion rate in urine (Weisel et al., 1999). The three sets of data were log-transformed to produce normally distributed data sets. The cut-point was arbitrarily set such that the subjects in the lowest tertile of exposure were classified as not exposed. Those in the higher two tertiles were then classified as exposed. The data sets were then compared, assuming that the urine excretion rate data represented the "true" exposure and the other two data sets represented the "observed" exposure. Sensitivity and specificity tables were set up by determining the change in exposure status for each subject (Tables 3-12 and 3-13).

Table 3-12. Sensitivity and specificity table of urine excretion rate vs. ingestion dose

	True exposure (urine excretion rate)		
Imperfect exposure (ingestion dose)	**Yes**	**No**	**Total**
Yes	15	3	18
No	16	13	29
Total	31	16	47

Sensitivity: 48%
Specificity: 81%
r^2 = 0.53 (for "true" exposure regressed on ingestion dose)

Table 3-13. Sensitivity and specificity table of urine excretion rate vs. water concentration

	True exposure (urine excretion rate)		
Imperfect exposure (water concentration)	**Yes**	**No**	**Total**
Yes	19	7	26
No	12	9	21
Total	31	16	47

Sensitivity: 61%
Specificity: 56%
r^2 = -0.04 (for "true" exposure regressed on water concentration)

It can be seen that these three different measures of exposure result in very different, and quite low sensitivities and specificities. Assume that a hypothetical study exists for which the "true" exposure measure of excretion rate results in an OR of 2.5. Applying this information to Table 3-16, it can be seen that if this hypothetical study used DBP concentrations in water as the measure of exposure rather than excretion dose, the true OR of 2.5, assuming that 30% of the non-diseased population is exposed, would be approximately 1.09. Similarly, if the same hypothetical study used ingestion dose as the measure of exposure instead of excretion rate, the true OR, assuming that 30% of the non-diseased population is exposed, would be approximately 1.35. It can be seen that the measure used for the exposure can have a significant effect on the strength of association of a study.

The data collected in City A and City B, as well as the monitoring data from City A, City B, and City C can be treated in a similar fashion to the exercise above. In reiterating this process with actual DBP concentration data, the relationships between different measures of exposure and their effect on the strength of association of a study can be investigated. As discussed in Chapter 1, many studies use quarterly monitoring data from the water treatment plants as the measure of the exposure of their study subjects. However, Chapter 1 also discussed the pitfalls of using these data as the measure of exposure for subjects in reproductive epidemiology studies.

With the data obtained from the water utilities, DBP concentrations at the plant can be compared with DBP concentrations in the distribution system, in locations that may be closer to the study subjects and therefore be more representative of the water concentrations to which the subjects will be exposed. This relationship can also be investigated between the water treatment plant and people's homes using the original data collected on a daily basis. The differences in attenuation of the OR between the monitoring data, which is taken monthly, and the original data, which was collected daily, will be investigated. These differences may suggest that more frequent measures should be employed in epidemiology studies.

THM concentrations have been used as a surrogate measure for other possible causal agents. The original data and monitoring data on the concentrations of many different DBPs can be employed in this method to determine what effect this aspect of exposure misclassification may have on the results of a study.

Treatment of the many different data sets available will provide an extensive picture of the effect the factors affecting water concentration and exposure to DBPs may have on the strength of association of these reproductive epidemiology studies.

Continuous treatment of data

The original and monitoring data can also be treated as continuous data in the investigation of the effects of misclassification on the results of epidemiological studies. Table 3-10 will be employed. The same data sets used in the categorical data treatment will be used in the continuous data treatment. Two data sets at a time can be compared by determining a regression coefficient. This regression coefficient or validity coefficient can then be applied to Table 3-10 and the resulting attenuation in a hypothetical OR can be determined. Using the exercise above as an example, the r^2 for urine excretion and ingestion dose is 0.53. Consulting Table 3-10, and again assuming a hypothetical "true" OR of 2.5, the resulting observed OR is seen to be between 1.58 and 1.73, assuming the OR is calculated per unit change.

This method can be used to investigate the same misclassification issues as were investigated treating the data as categorical. In addition, a comparison can be made of the effect of treating these data sets as categorical *vs*. continuous.

These tables provide the tools necessary to investigate how different measures of exposure can affect the measure of strength of association of epidemiologic studies. The original and historical data sets will be used for this purpose. Correlation coefficients, sensitivity, and specificity can be calculated from selected data sets and compared with the values in Tables 3-6 and 3-10. Using these methods, the effect of exposure misclassification on the measure of strength of exposure will be seen in quantitative terms, and we can begin to understand how "good" exposure data must be to provide close estimates of the "true" OR of a study.

In this chapter, exposure misclassification was discussed along with the effect of exposure misclassification on the OR of epidemiology studies. Methods for quantifying the effect of exposure misclassification on OR attenuation were introduced. A plan for applying these methods to real data was developed.

References

Armstrong, Ben G., 1998. Effect of Measurement Error on Epidemiological Studies of Environmental and Occupationsl Exposures. *Occupational and Environmental Medicine* **55**: 651-656.

Armstrong, Bruce K., Emily White and Rodolfo Saracci, 1992. *Principles of Exposure Measurement in Epidemiology*. Oxford, Oxford University Press.

Bove, F.J., M.C. Fulcomer, J.B. Koltz, J. Esmart, E.M. Dufficy and R.T. Zagraniski, 1992. *Report on Phase IV-A: Public Drinking Water Contamination and Birthweight Fetal Deaths, and Birth Defects, A Cross Sectional Study*. New Jersey Department of Health. 1992.

de Klerk, Nicholas H., Dallas R. English and Bruce K. Armstrong, 1989. A Review of the Effects of Random Measurement Error on Relative Risk Estimates in Epidemiological Studies. *International Journal of Epidemiology* **18**(3): 705-712.

Delpizzo, Vincent and Joanna L. Borghesi, 1995. Exposure Measurement Errors, Risk Estimate and Statistical Power in Case-Control Studies Using Dichotomous Analysis of a Continuous Exposure Variable. *International Journal of Epidemiology* **24**(4): 851-862.

Dosemeci, Mustafa, Sholom Wacholder and Jay H. Lubin, 1990. Does Nondifferential Misclassification of Exposure Always Bias a True Effect Toward the Null Value? *American Journal of Epidemiology* **132**(4): 746-748.

Flegal, Katherine M., Penelope M. Keyl and F. Javier Nieto, 1991. Differential Misclassification Arising from Nondifferential Errors in Exposure Measurement. *American Journal of Epidemiology* **134**(10): 1233-1244.

Hennekens, Charles H. and Julie E. Buring, 1987. *Epidemiology in Medicine*. Boston, Little, Brown and company.

Kelsey, Jennifer L., W. D. Thompson and Alfred S. Evans, 1986. *Methods in Observational Epidemiology*. New York, Oxford University Press.

Klotz, Judith B. and Laurie A. Pyrch, 1999. Neural Tube Defects and Drinking Water Disinfection By-Products. *Epidemiology* **10**(4): 383-390.

Reif, John S., Annette Bachand and Melvin Andersen, 2000. *Reproductive and Developmental Effects of Disinfection By-Products*. Ottawa, Ontario. Health Canada. October 31, 2000.

Rothman, Kenneth J. and Sander Greenland, 1998. *Modern Epidemiology*. Philadelphia, Lippincott-Raven Publishers.

Wacholder, Sholom, Mustafa Dosmeci and Jay H. Lubin, 1991. Blind Assignment of Exposure Does Not Always Prevent Differential Misclassification. *American Journal of Epidemiology* **134**(4): 433-437.

Weisel, Clifford P., Hekap Kim, Patricia Haltmeier and Judith B. Klotz, 1999. Exposure Estimates to Disinfection By-Products of Chlorinated Drinking Water. *Environmental Health Perspectives* **107**(2): 103-110.

Wu, Marion L., Alice S. Whittemore and Dexter L. Jung, 1986. Errors in Reported Dietary Intakes. I. Short Term Recall. *American Journal of Epidemiology* **124**(5): 826-835.

Chapter 4

Data Sources and Data Collection

Monitoring data

Monitoring data was generously provided by the water utilities in three Canadian cities: City A, City B and City C. The data provided constituted regular monitoring data in each of the three cities. A summary of the data provided by the water utilities is presented in Tables 4-1 to 4-3.

Table 4-1. City A monitoring data

Compounds or parameters	Frequency of sampling	Dates	Location
TTHM, TCM, BDCM, CDBM, TBM	Daily	1/1/97-31/12/00	WTP2, WTP1 res.
pH, turbidity, temperature, VOC, colour, hardness, alkalinity, chlorine dose, alum, ammonia	Daily	01/03/97-31/12/00	WTP2, WTP1 raw
TOC	Weekly	01/01/97-31/12/00	WTP2 raw and res. WTP1 raw and res.
TTHM, TCM, BDCM, CDBM, TBM	Several times/month	01/09/97-31/12/00	Random DS
TTHM	Monthly	01/01/97-31/12/99	WTP2, WTP1 res. DS
HAA_7, THAA	Monthly	01/02/96-31/12/00	WTP2, WTP1 res.
THAN, THK, chloropicrin, CH	Monthly	01/01/97-31/12/00	WTP2, WTP1 res.

WTP2: Water treatment plant #2
WTP1: Water treatment plant #1
DS: Distribution System
Res.: Reservoir
Raw: untreated water

Table 4-2. City B monitoring data

Compounds or parameters	Frequency of sampling	Dates	Location
Chlorine dose	Daily	01/01/00-31/12/00	WTP1
Water temperature	Daily	01/06/99-30/11/00	WTP1 raw
pH	Weekly	01/09/00-30-11/00	DS
TTHM, TCM, BDCM, CBDM, TBM	Monthly	01/10/98-31/12/00	WTP1, WTP2 eff. DS
THAA, DCAA, TCAA, BCAA, DBAA, MBAA	Monthly	01/10/98-31/12/00	WTP1, WTP2 eff. DS
TCAN, DCAN, BCAN, DBAN	Monthly	01/10/98-31/12/00	WTP1, WTP2 eff. DS
TOC, Total solvent-extractable DBPs, CH, DCA, TCNM	Monthly	01/10/98-31/12/00	WTP1, WTP2 eff. DS

WTP1: Water treatment plant #1
WTP2: Water treatment plant #2
DS: Distribution System
Eff.: Effluent

Table 4-3. City C monitoring data

Compounds or parameters	Frequency of sampling	Dates	Location
TTHM, TCM, BDCM, CDBM, TBM	Monthly	01/01/99-30/11/00	Intake through DS
THAA and HAA_6	Monthly	01/01/99-30/11/00	Intake through DS
TAld, Ald_8	Monthly	01/01/99-31/12/99	Intake through DS
DCAA, TCAA	Monthly	01/01/99-31/12/00	Intake through DS
TTHM	Monthly	01/01/94-31/12/98	Intake through DS
THAA	Monthly	01/01/97-31/12/98	Intake through DS
TAld	Monthly	01/01/98-31/12/98	Intake through DS

Description of water utilities

City A, City B and City C were chosen as the data source locations because of the effect certain differences in water treatment and distribution at each location may have on the resulting DBP concentrations both in the water treatment plants and in the distribution systems. Additionally, these water utilities have had DBP monitoring programs in place for long enough to provide good historical data sets.

City A

City A has two water treatment plants, water treatment plant #1 and water treatment plant #2, that treat water from the same river. Both plants employ conventional treatment, including coagulation, flocculation, sedimentation, and filters. The filters at water treatment plant #2 are dual media anthracite/sand filters, while those at water treatment plant #1 during the time the data were collected were sand filters. Water treatment plant #1 has two plants, built in 1947 and 1956, while water treatment plant #2 was built in the mid-1970s. Primary disinfection in City A is by chlorination, with the residual provided by chloramine. Free chlorine is added before filtration to allow for a free chlorine contact time. After the filters, ammonia is added to react with the chlorine and provide a stable combined chlorine residual. The distribution system in City A is open and served by both water treatment plants. The water at any particular location at any time could be from one plant exclusively or from a mix of the two, depending on demand, which pumps are employed, and whether the water has passed through one of several reservoirs in the distribution system. Sampling points for the daily data were located throughout the system since the source plant for any point in the system can change on an hourly basis. Sampling locations were chosen based on whether one of the following conditions were likely to occur the majority of the time: mostly water treatment plant #1 water (very close to water treatment plant #1), mostly water treatment plant #2 water (very close to water treatment plant #2), mixed water (short residence time), mixed water (long residence time). The locations were chosen to provide as a wide range of residence times and source as possible within the confines of the study parameters. Monitoring locations for the monthly historical data are more numerous and cover the entire distribution system.

City B

City B also has two water treatment plants, water treatment plant #1 and water treatment plant #2. Water treatment plant #1's source water is a reservoir, which is fed by a river. Water treatment plant #2 also takes its water from a reservoir fed by a river. Both plants also use conventional treatment with coagulation, flocculation, sedimentation, and filtration. Both plants employ dual media anthracite/sand filters. Free chlorine is added twice: to the raw water before treatment, and to the treated water just before entry into the distribution system. No ammonia is added because free chlorine, not chloramine is used to maintain residual disinfection. The distribution system in City B is also open; however, it is divided into pressure zones. The pressure zones can be identified as being supplied by either water treatment plant #1 or water treatment plant #2 water treatment plants, or both. The sampling locations for the daily data were located exclusively in the water treatment plant #1 zone. Water treatment plant #1 water was chosen because the residence time of the raw water in the reservoir before treatment increases the NOM, resulting in higher DBP concentrations. This is in contrast to the relatively lower DBP concentrations produced by water treatment plant #2 or the plants in City A. In addition, the lab where sample storage, preparation, and analysis occurred was located at water treatment plant #1. The historical monitoring data examined in the following chapters is exclusively from the water treatment plant #1 plant and sampling points in the water treatment plant #1 zone to enable comparisons between the monthly historical data and the daily data.

City C
City C has a unique system in which water is not treated in one particular water treatment plant, but rather chlorinated at several points in the distribution system. Water is piped in via an aqueduct from a lake 160 km from the city. The residence time in this aqueduct is approximately 30 hours. The water is first chlorinated at the intake point in the lake and large suspended items are filtered out. From the aqueduct, the water is held in a large open-air reservoir near the city from which the water is distributed along two smaller aqueducts to smaller reservoirs and, ultimately, to the whole distribution system. Residence time in the large reservoir is between 48 hours and 10 days. Chlorine is added as the water exits the large reservoir and again as the water exits the smaller reservoirs. Free chlorine is the disinfectant used in City C. DBP concentrations are expected to be high in the distribution system due to the high NOM in the source water, particularly in the summer, as well as the long retention times from the lake to the consumer, during which the chlorine has the opportunity to react with the NOM.

Experimental Data

Daily sampling for DBPs in drinking water

Sample Bottle Preparation: City A
250 mL high-density polyethylene bottles (Wheaton) were prepared with 20 drops of 250 g/L NH_4Cl as dechlorination agent, following U.S. EPA methods 551.1 and 552.2 (U.S. EPA, 1995a; U.S. EPA, 1995b). Due to the organic nature of the bottles, the bottles were tested before sampling for addition of DBPs or escape of volatile DBPs. No additive or subtractive effects were found for the time period that samples would be in the bottles before extraction and analysis. Samples for the analysis of HAA and THM/HAN were drawn from the same bottles

Sample Bottle Preparation: City B
40 mL open top clear glass vials with Teflon-lined septa, delivered as "precleaned and certified", but without documentation (VWR) were used for HAA and THM samples. HAA bottles were prepared with 3 drops of 250 g/L NH_4CL as dechlorination agent while THM bottles were prepared with 2 drops of 80g/L sodium thiosulphate as dechlorination agent, according to City B analysis laboratory sampling procedures.

Sampling method: City A and City B
Samples of cold tap water were taken daily at approximately the same time each day for 29 days. The tap water was flushed until cold (approx. 2-3 minutes) to ensure sampled water was not affected by an extended residence time in the

household plumbing. Bottles were filled to the top, with no headspace, being careful not to overflow the sample and lose part of the dechlorination agent.

Sample handling: City A and City B

All participants in the study were given coolers and ice packs with which to transport their samples to the collection points daily (the workplace in both cities). On weekends, samples were stored in the participants' fridges until they could be brought to work in the coolers. In this way, refrigeration of the samples at all times was ensured. City B samples were analysed in two batches, therefore samples were stored in a refrigerator in the City B analysis laboratories until sample preparation and analysis could be done. City A samples were analysed on a daily basis and samples left in the bottles were stored in a refrigerator or refrigerated cold room. Duplicate field samples were taken each day on a rotating basis based on location. Duplicate field samples were sampled, handled and analysed exactly as the samples. One field blank was analyzed from each location. The field blanks were sampling vials filled with Milli-Q water and the appropriate quenching agent. The samples were transported to the sampling locations, where they were opened for the length of time of sampling, closed, and handled and analysed just as normal samples.

Sample Preparation: City B

HAAs

A 13.2 mL aliquot was withdrawn from each 40 mL sample vial using an adjustable pipettor. 6μL of 2-bromopropionic acid surrogate standard was added. 1.5 mL concentrated H_2SO_4, 3g $CuSO_4 \cdot 5H_2O$, 12g Na_2SO_4 and 3 mL of MTBE extraction solvent containing 1, 2, 3-trichloropropane internal standard were added to the sample. The samples were shaken mechanically for 9 minutes, after which the phases were allowed to separate for 3 minutes. The MTBE layer was transferred to 13x100 mm test tubes using Pasteur pipettes at which time 2 mL of 10% H_2SO_4/methanol solution was added. The capped test tubes were placed in a pre-warmed heating block at 50°C for 1 hour. The test tubes were then removed from the heating block and allowed to cool. 8 mL of saturated $NaHCO_3$ solution was added 1 mL at a time, mixing gently between additions. The test tube was then vortexed for 1 minute. The two phases were allowed to separate, after which the upper MTBE layer was transferred to a 2 mL autosampler vial.

THM

THM samples in their sampling vials were placed directly onto the autosampler of the purge-and-trap.

Sample Preparation: City A

HAAs

40 mL of the water sample was measured into a 50 mL centrifuge tube (Corning Inc.) using a graduated cylinder. 2 mL of concentrated H_2SO_4, 12g of Na_2SO_4, and 4 mL of MTBE were added in succession to the water sample. The tube was capped and the sample was shaken on a Janke and Kunkel Type VX8 IKA-

VIBRAX-VXR shaker for 5 minutes at 1300 rpm. The two phases were allowed to separate and the MTBE layer was drawn off using a Pasteur pipette and transferred to a 16 mL glass vial with Teflon-coated cap (Fisher). 2 mL of acidified methanol (10% solution of H_2SO_4 in methanol) was added to the extract and the vials were placed on a 50°C heating block for 1 hour for methylation to take place. The extract mixture was then allowed to cool before 8 mL of a saturated mixture of $NaHCO_3$ in Milli-Q water was added dropwise to neutralize the mixture. The organic layer was then drawn off into a 2 mL autosampler vial (Supelco) using a Pasteur pipette and capped with a Teflon-coated crimp cap (Supelco). One sample per analysis day was split into duplicate samples. These lab duplicates were prepared and analysed exactly as other samples. A Milli-Q water blank was prepared and analysed on each sample day.

THMs

40 mL of the water sample was measured into a 50 mL centrifuge tube (Corning Inc.) using a graduated cylinder. 12 g of sodium sulphate and 4 mL of MTBE were added to the tube. The tube was capped and shaken for 5 minutes using a Janke and Kunkel Type VX8 IKA-VIBRAX-VXR shaker for 5 minutes at 1300 rpm. The two phases were allowed to separate and the MTBE layer was drawn off using a Pasteur pipette. The MTBE layer was transferred to a 2 mL autosampler vial (Supelco) and capped with a Teflon-coated crimp cap (Supelco). One sample per analysis day was split into duplicate samples. These lab duplicates were prepared and analysed exactly as other samples. A Milli-Q water blank was prepared and analysed on each sample day.

Sample analysis: City A

Both the HAA and volatiles samples were analyzed using gas chromatography with electron capture detection (GC/ECD) on an Agilent 6890 GC/ECD using a J&W Scientific DB-5.625 column: 30 m x 0.25 mm (i.d.) x 0.25 μm. The detector temperature was 300 °C and the injector temperature was 200 °C. Initial oven temperature was held at 35 °C for 15 minutes, then ramped at 10 °C/min to 100 °C and held for 5 minutes. The temperature was then ramped to 120 °C at 10 °C/min and held for 5 minutes. Finally, the temperature was ramped at 20 °C/min to 200 °C and held for 1 minute. The average argon-methane 5% makeup gas flow was 60 mL/min. The average helium carrier gas flow velocity was 9.3 psi. 2 μL of each sample was injected at a 5:1 split ratio. A solvent blank was injected at several times during each analysis run.

Sample analysis: City B

The HAA samples were analyzed using gas chromatography with electron capture detection (GC/ECD) on a Hewlett Packard (HP) 5890 Series II GC/ECD using a J&W Scientific DB-5 column: 30 m x 0.25 mm (i.d.) x 1 μm. The detector temperature was 300 °C and the injector temperature was 200 °C. 1 μL of sample was injected splitless onto the column. Initial oven temperature was held at 40 °C for 8 minutes, then ramped at 5 °C/min to 120 °C. The temperature was then ramped to 210 °C at 15 °C/min. The average helium carrier gas flow velocity was

1.28 mL/min. THM samples were placed directly onto the purge and trap autosampler, which was attached to a Hewlett Packard (HP) 5890 Series II GC/ECD. 10 mL of the sample were extracted and purged for 6 minutes with He. The sample was desorbed by raising the temperature of the trap from ambient to 245°C over a one minute period. The column used was a DB-VRX column: 30m x 0.32 mm (i.d.) x 1.8 µm. The initial oven temperature was 35°C and held for 1 minute after which the temperature was ramped at 5°C/min to a final temperature of 180°C. The ECD flow was approximately 40 mL/min with He carrier and N_2 auxiliary gases.

Summary statistics for experimental and monitoring data

The data are summarized in Tables 4-4 to 4-8. The City A and City B daily water treatment plant data are fairly comparable in terms of concentrations, although the variation in City A is slightly higher than in City B. The distribution system concentrations in City B are higher than in City A, as expected from a chlorinated system (City B) vs. a chloraminated system (City A). City B monthly data show higher concentrations and greater variation at both the water treatment plant and the distribution system than the City A monthly data. City C is a unique case of a chlorinated system, with the inclusion of an open reservoir between the first and second chlorination points (sampling points #2 and #3). There is a distinct decrease in all DBP concentrations between these two points that is attributed to residence time in the open reservoir. After the addition of chlorine at sampling point #3, the DBP concentrations rise again. Concentrations of all DBPs are substantially higher in City C at all points than in either City A or City B. Probable reasons for the higher DBP concentrations include the long chlorine contact time both before and after the open reservoir, and high organic content in the raw water from the source lake. In addition, conventional treatment that would further remove natural organic matter before chlorination is not applied.

Table 4-4. City A daily summary statistics

WTP1 (n=27)	TCM	DCAN	DCAA	TCAA
Mean (µg/L)	16	1.3	8.2	11
Median (µg/L)	16	1.2	7.9	11
Std. Dev. (µg/L)	5.2	0.49	4.2	3.1
Rel. % Std. Dev.	32	37	51	28
Max. (µg/L)	26	2.4	16	16
Min. (µg/L)	6.8	0.74	2.9	5.4
WTP2 (n=27)	**TCM**	**DCAN**	**DCAA**	**TCAA**
Mean (µg/L)	16	1.2	7.4	11
Median (µg/L)	17	1.3	8.5	10
Std. Dev. (µg/L)	5.4	0.39	3.4	3.4
Rel. % Std. Dev.	33	33	46	32
Max. (µg/L)	25	1.9	14	18
Min. (µg/L)	6.8	0.74	2.3	5.1
A01 (n=29)	**TCM**	**DCAN**	**DCAA**	**TCAA**
Mean (µg/L)	16	1.3	8.7	12
Median (µg/L)	16	1.4	8.0	12
Std. Dev. (µg/L)	4.2	0.42	4.1	3.1
Rel. % Std. Dev.	26	33	47	25
Max. (µg/L)	26	2.2	16	17
Min. (µg/L)	8.1	0.74	3.1	6.7
A02 (n=29)	**TCM**	**DCAN**	**DCAA**	**TCAA**
Mean (µg/L)	16	1.3	8.4	11
Median (µg/L)	15	1.2	7.1	11
Std. Dev. (µg/L)	4.0	0.38	4.2	3.0
Rel. % Std. Dev.	25	29	50	27
Max. (µg/L)	28	2.3	17	17
Min. (µg/L)	8.9	0.91	2.3	5.2
A03 (n=29)	**TCM**	**DCAN**	**DCAA**	**TCAA**
Mean (µg/L)	17	1.5	10	12
Median (µg/L)	17	1.5	10	12
Std. Dev. (µg/L)	3.8	0.40	4.6	2.9
Rel. % Std. Dev.	22	27	47	25
Max. (µg/L)	23	2.3	18	17
Min. (µg/L)	8.6	0.81	2.6	5.2
A04 (n=29)	**TCM**	**DCAN**	**DCAA**	**TCAA**
Mean (µg/L)	17	1.4	8.9	11
Median (µg/L)	15	1.3	8.6	11
Std. Dev. (µg/L)	5.5	0.47	4.8	3.3
Rel. % Std. Dev.	33	33	54	29
Max. (µg/L)	28	2.3	19	17
Min. (µg/L)	6.5	0.77	2.6	5.4

Table 4-5. City B daily summary statistics

WTP (n=29)	TCM	BDCM	DCAA	TCAA
Mean (μg/L)	19	1.0	9.8	9.8
Median (μg/L)	19	1.0	10	10
Std. Dev. (μg/L)	2.0	0.07	1.0	1.4
Rel. % Std. Dev.	10	6.8	10	14
Max. (μg/L)	23	1.2	12	12
Min. (μg/L)	13	0.85	7.4	7.4
B01 (n=28)	**TCM**	**BDCM**	**DCAA**	**TCAA**
Mean (μg/L)	24	1.3	11	13
Median (μg/L)	25	1.4	11	13
Std. Dev. (μg/L)	4.2	0.18	1.9	2.6
Rel. % Std. Dev.	17	13	17	20
Max. (μg/L)	30	1.8	15	18
Min. (μg/L)	15	0.99	8.5	8.5
B02 (n=29)	**TCM**	**BDCM**	**DCAA**	**TCAA**
Mean (μg/L)	30	1.6	15	19
Median (μg/L)	30	1.6	15	19
Std. Dev. (μg/L)	2.4	0.13	1.0	3.2
Rel. % Std. Dev.	8.0	7.9	6.7	16
Max. (μg/L)	39	1.9	18	25
Min. (μg/L)	26	1.3	13	13
B03 (n=29)	**TCM**	**BDCM**	**DCAA**	**TCAA**
Mean (μg/L)	30	1.5	13	17
Median (μg/L)	29	1.5	14	17
Std. Dev. (μg/L)	4.1	0.24	1.8	3.3
Rel. % Std. Dev.	14	16	13	20
Max. (μg/L)	39	2.1	17	26
Min. (μg/L)	23	1.2	9.1	11.3
B04 (n=28)	**TCM**	**BDCM**	**DCAA**	**TCAA**
Mean (μg/L)	33	1.8	12	24
Median (μg/L)	33	1.8	12	23
Std. Dev. (μg/L)	2.0	0.07	1.5	6.4
Rel. % Std. Dev.	6.1	4.2	13	27
Max. (μg/L)	38	1.9	15	49
Min. (μg/L)	29	1.6	8.6	16

Table 4-6. City A monthly summary statistics

	WTP1 DCAA	WTP1 TCAA	WTP1 THAA	WTP1 TCM	WTP1 BDCM	WTP1 TTHM
Mean (μg/L)	5.9	3.7	9.3	11	0.66	7.7
Median (μg/L)	4.8	2.9	8.0	8.2	0.64	5.1
Std. Dev. (μg/L)	3.8	3.1	6.5	7.2	0.21	7.1
Rel. % Std. Dev.	64	84	70	69	31	92
Max. (μg/L)	15	12	27	29	1.3	27
Min. (μg/L)	0.70	0.20	1.0	1.3	0.40	1.0
Sample size (n)	49	44	48	48	48	84

	WTP2 DCAA	WTP2 TCAA	WTP2 THAA	WTP2 TCM	WTP2 BDCM	WTP2 TTHM
Mean (μg/L)	5.6	3.7	9.2	11.6	0.67	9.5
Median (μg/L)	4.8	2.6	6.8	7.5	0.59	6.4
Std. Dev. (μg/L)	3.7	3.4	6.9	8.3	0.19	7.9
Rel. % Std. Dev.	66	91	75	72	28	84
Max. (μg/L)	16	15	31	31	1.3	31
Min. (μg/L)	0.60	0.20	0.90	2.3	0.41	1.0
Sample size (n)	48	43	48	48	48	84

	TTHM1	TTHM2	TTHM3	TTHM4	TTHM5	TTHM6
Mean (μg/L)	7.1	9.2	9.3	8.9	8.7	7.3
Median (μg/L)	4.0	6.5	7.0	6.0	7.0	5.0
Std. Dev. (μg/L)	6.5	8.6	8.5	8.3	7.5	6.9
Rel. % Std. Dev.	91	94	92	93	85	94
Max. (μg/L)	27	32	32	31	33	32
Min. (μg/L)	1.0	1.0	1.0	1.0	1.0	1.0
Sample size (n)	47	50	50	47	45	50

	TTHM7	TTHM8	TTHM9	TTHM10	TTHM11	TTHM12
Mean (μg/L)	6.8	8.2	9.5	9.4	9.8	6.5
Median (μg/L)	4.0	6.0	6.5	6	6.5	4
Std. Dev. (μg/L)	5.9	8.4	9.1	8.9	9.3	6.9
Rel. % Std. Dev.	87	100	96	95	95	110
Max. (μg/L)	24	34	36	33	36	30
Min. (μg/L)	1.0	1.0	1.0	1.0	1.0	1.0
Sample size (n)	50	45	48	50	50	50

Table 4-7. City B monthly summary statistics

	WTP TCM	WTP BDCM	WTP TTHM	WTP DCAA	WTP TCAA	WTP THAA
Mean (μg/L)	17	1.1	15	9.1	7.9	17
Median (μg/L)	13	1.0	13	7.9	7.9	16
Std. Dev. (μg/L)	13	0.53	10	6.2	5.5	12
Rel. % Std. Dev.	79	49	69	68	69	68
Max. (μg/L)	51	2.5	51	27	22	49
Min. (μg/L)	3.9	0.34	4.3	2.6	1.9	5.1
Sample size (n)	36	36	27	21	21	22
	B01TCM	**B01BDCM**	**B01TTHM**	**B01DCAA**	**B01TCAA**	**B01THAA**
Mean (μg/L)	22	1.4	23	11	13	24
Median (μg/L)	22	1.4	22	8.0	11	20
Std. Dev. (μg/L)	11	0.38	12	7.5	7.7	15
Rel. % Std. Dev.	50	27	54	66	60	62
Max. (μg/L)	52	1.9	54	32	32	64
Min. (μg/L)	7.1	0.79	8.3	3.9	4.7	9.6
Sample size (n)	17	17	27	25	25	26
	B02TCM	**B02BDCM**	**B02TTHM**	**B02DCAA**	**B02TCAA**	**B02THAA**
Mean (μg/L)	25	1.5	26	13	14	28
Median (μg/L)	26	1.6	25	11	14	26
Std. Dev. (μg/L)	13	0.41	14	8.3	9.5	17
Rel. % Std. Dev.	54	27	54	64	66	61
Max. (μg/L)	61	2.2	63	36	38	75
Min. (μg/L)	7.5	0.91	8.7	4.5	0.92	11
Sample size (n)	17	17	26	25	25	26
	B03TCM	**B03BDCM**	**B03TTHM**	**B03DCAA**	**B03TCAA**	**B03THAA**
Mean (μg/L)	27	1.6	28	12	16	27
Median (μg/L)	25	1.5	25	9.6	14	22
Std. Dev. (μg/L)	16	0.44	16	8.8	11	19
Rel. % Std. Dev.	58	27	58	70	68	71
Max. (μg/L)	74	2.5	76	41	50	91
Min. (μg/L)	7.5	0.92	8.7	4.4	5.2	11
Sample size (n)	17	17	26	25	25	26
	B04TCM	**B04BDCM**	**B04TTHM**	**B04DCAA**	**B04TCAA**	**B04THAA**
Mean (μg/L)	21	1.3	22	11	12	23
Median (μg/L)	20	1.4	19	7.4	10.8	18
Std. Dev. (μg/L)	11	0.33	11	7.5	7.2	14
Rel. % Std. Dev.	52	25	52	69	62	64
Max. (μg/L)	52	1.8	54	31	32	63
Min. (μg/L)	6.7	0.78	7.8	3.3	3.9	8.4
Sample size (n)	17	17	25	24	24	25

Table 4-8. City C monthly summary statistics

	THAA2	DCAA2	TCAA2	TTHM2	TCM2	BDCM2
Mean (µg/L)	79	25	52	102	100	9.0
Median (µg/L)	75	26	50	98	93	9.0
Std. Dev. (µg/L)	29	12	19	33	33	2.9
Rel. % Std. Dev.	36	45	36	33	33	32
Max. (µg/L)	150	47	92	190	170	14
Min. (µg/L)	25	1.3	24	26	25	1.4
Sample size (n)	35	19	19	68	21	21
	THAA3	**DCAA3**	**TCAA3**	**TTHM3**	**TCM3**	**BDCM3**
Mean (µg/L)	44	8.2	39	51	50	4.6
Median (µg/L)	48	2.6	46	43	32	4.6
Std. Dev. (µg/L)	23	11	18	30	36	3.0
Rel. % Std. Dev.	51	140	47	59	72	64
Max. (µg/L)	83	29	60	160	89	8.0
Min. (µg/L)	3.8	1.1	6.0	1.4	1.4	1.3
Sample size (n)	24	6	7	56	7	6
	THAA4	**DCAA4**	**TCAA4**	**TTHM4**	**TCM4**	**BDCM4**
Mean (µg/L)	54	5.3	44	55	55	4.8
Median (µg/L)	53	4.9	42	50	47	2.4
Std. Dev. (µg/L)	18	3.7	12	33	32	3.7
Rel. % Std. Dev.	33	70	26	59	59	77
Max. (µg/L)	100	14	72	130	120	12
Min. (µg/L)	11	1.8	30	3.1	23	1.4
Sample size (n)	28	10	14	31	17	17
	THAA5	**DCAA5**	**TCAA5**	**TTHM5**	**TCM5**	**BDCM5**
Mean (µg/L)	72	14	50	80	79	8.1
Median (µg/L)	68	13	48	80	73	7.1
Std. Dev. (µg/L)	26	6.3	16	25	24	2.8
Rel. % Std. Dev.	36	45	33	31	30	35
Max. (µg/L)	130	33	95	160	120	13
Min. (µg/L)	16	6.2	16	13	39	4.1
Sample size (n)	38	17	18	73	20	20
	THAA6	**DCAA6**	**TCAA6**	**TTHM6**	**TCM6**	**BDCM6**
Mean (µg/L)	73	14	56	81	72	7.3
Median	74	14	60	78	69	6.7
Std. Dev.	24	7.1	19	24	23	2.6
Rel. % Std. Dev.	32	51	33	30	32	36
Max. (µg/L)	130	37	92	130	120	12
Min. (µg/L)	18	0.8	17	38	36	3.8
Sample size (n)	38	20	20	39	22	22

Table 4-8. Continued City C monthly summary statistics

	THAA7	DCAA7	TCAA7	TTHM7	TCM7	BDCM7
Mean (μg/L)	81	18	59	96	94	9.4
Median (μg/L)	81	17	59	92	83	9.0
Std. Dev. (μg/L)	22	9.5	16	25	26	2.6
Rel. % Std. Dev.	27	52	27	26	28	28
Max. (μg/L)	140	33	98	150	140	14
Min. (μg/L)	28	2.0	26	46	56	5.8
Sample size (n)	45	21	21	80	23	23

	THAA8	DCAA8	TCAA8	TTHM8	TCM8	BDCM8
Mean (μg/L)	86	20	60	92	91	9.2
Median (μg/L)	76	18	58	92	90	8.9
Std. Dev. (μg/L)	25	6.6	15	21	22	2.5
Rel. % Std. Dev.	29	34	25	23	25	27
Max. (μg/L)	160	34	100	140	120	14
Min. (μg/L)	48	11	42	47	58	5.8
Sample size (n)	45	21	21	79	23	23

	THAA9	DCAA9	TCAA9	TTHM9	TCM9	BDCM9
Mean (μg/L)	79	18	56	93	92	9.7
Median (μg/L)	76	18	52	90	84	8.5
Std. Dev. (μg/L)	21	6.3	15	22	24	3.8
Rel. % Std. Dev.	27	35	26	24	26	39
Max. (μg/L)	140	29	95	150	130	23
Min. (μg/L)	42	3.2	38	44	58	5.9
Sample size (n)	46	22	22	81	24	24

	THAA10	DCAA10	TCAA10	TTHM10	TCM10	BDCM10
Mean (μg/L)	88	22	61	100	99	9.8
Median (μg/L)	84	20	59	97	91	9.6
Std. Dev. (μg/L)	24	8.0	14	24	24	2.5
Rel. % Std. Dev.	28	36	22	24	24	26
Max. (μg/L)	160	41	95	170	140	14
Min. (μg/L)	43	12	43	52	64	6.5
Sample size (n)	45	22	22	81	24	24

	THAA11	DCAA11	TCAA11	TTHM11	TCM11	BDCM11
Mean (μg/L)	88	23	64	110	101	9.8
Median (μg/L)	85	22	63	110	104	9.4
Std. Dev. (μg/L)	20	6.4	15	27	24	2.7
Rel. % Std. Dev.	23	27	23	25	23	28
Max. (μg/L)	140	37	100	160	150	14
Min. (μg/L)	50	12	43	61	62	5.6
Sample size (n)	24	22	22	26	24	24

Table 4-8. Continued City C monthly summary statistics

	THAA12	DCAA12	TCAA12	TTHM12	TCM12	BDCM12
Mean (µg/L)	79	21	57	110	99	9.9
Median (µg/L)	76	20	53	97	88	9.2
Std. Dev. (µg/L)	24	7.6	14	29	25	2.7
Rel. % Std. Dev.	31	36	25	27	26	27
Max. (µg/L)	160	37	102	160	140	15
Min. (µg/L)	40	9.1	38	50	59	6.3
Sample size (n)	23	21	21	26	24	24

	THAA13	DCAA13	TCAA13	TTHM13	TCM13	BDCM13
Mean (µg/L)	70	16	54	110	100	10
Median (µg/L)	72	16	53	110	97	9.3
Std. Dev. (µg/L)	16	5.6	11	25	23	2.4
Rel. % Std. Dev.	23	35	21	23	22	24
Max. (µg/L)	110	24	79	150	140	14
Min. (µg/L)	35	3.8	31	67	70	6.6
Sample size (n)	23	22	22	26	24	24

	THAA14	DCAA14	TCAA14	TTHM14	TCM14	BDCM14
Mean (µg/L)	87	24	62	110	105	10
Median (µg/L)	87	22	60	110	102	9.9
Std. Dev. (µg/L)	25	6.8	17	30	26	2.8
Rel. % Std. Dev.	29	29	27	27	25	27
Max. (µg/L)	160	37	100	180	170	17
Min. (µg/L)	46	12	41	53	65	6.4
Sample size (n)	23	21	21	26	24	24

	THAA15	DCAA15	TCAA15	TTHM15	TCM15	BDCM15
Mean (µg/L)	88	21	60	100	100	10
Median (µg/L)	84	21	60	100	100	9.7
Std. Dev. (µg/L)	24	7.8	16	22	25	2.5
Rel. % Std. Dev.	28	37	26	21	24	25
Max. (µg/L)	160	37	110	160	140	15
Min. (µg/L)	41	4.0	35	54	68	6.6
Sample size (n)	45	22	22	81	23	23

K_{ow} experiments

A series of experiments was done early in this research program to determine the octanol-water partition coefficients (K_{ow}) of DCAN and CH. The experimental method and final results are presented here. This section is not discussed in relation to exposure misclassification, although the results could be used in future research on exposure pathways.

As seen in Chapter 1, the octanol-water partition coefficient (K_{ow}) of DBPs can be used to predict the uptake and distribution of these chemicals in the body. The K_{ows} for several DBPs have been experimentally determined (Table 4.9). However, for other DBPs, namely the HANs and HAAs, the K_{ows} have been only estimated from computer programs such as Kowwin and MMPro. To further an understanding of the behaviour of some of these compounds, the K_{ows} for DCAN and CH were determined experimentally. The experimental method is presented below and the results are presented in Table 4-9.

Experimental method

The slow-stirring method was used to determine the octanol-water partition coefficient. In the slow-stirring method, the water and octanol phases are stirred slowly until equilibration is reached (normally between 2 and 8 days). Slow stirring and careful temperature control are necessary to prevent emulsion of the two phases (van Haelst et al., 1994). The experimental flasks were set up as described by Brooke and colleagues (Brooke et al., 1986). The flasks were submerged in a circulating water bath with the temperature controlled to 25 °C ± 0.5 °C. 900 mL of Millipore water and 100 mL of octanol were added to each flask. These phases were stirred overnight in the constant temperature water bath, allowing the phases to become mutually saturated. The water phase was then spiked with the solute to yield a concentration of 50 μg/L. The spiked flasks were stirred for 7 days to equilibrate and sampled on the seventh day.

Sample preparation

Samples for analysis from the water phase were prepared in accordance with EPA method 551.1 (U.S. EPA, 1995a), a liquid-liquid extraction method. A 30 mL aliquot of water sample was withdrawn from the flasks, salted with sodium sulphate, and shaken vigorously with 3 mL methyl-*tert*-butyl ether (MTBE). After separation of the phases, 100 μL of the MTBE layer was transferred to autosampler vials and diluted with 900 μL of MTBE spiked with internal standard. The octanol samples required no extraction prior to analysis. 970 μL of octanol sample was transferred directly into an autosampler vial with 30 μL of internal standard solution.

Sample analysis

The samples were analyzed using gas chromatography with electron capture detection (GC/ECD) on an Agilent 6890 GC/ECD using a J&W DB-5.625 column: 30 m x 0.25 mm (i.d.) x 0.25 μm. The detector temperature was 300 °C and the injector temperature was 205 °C. Initial oven temperature was 35 °C, then ramped at 20 °C/min to 220 °C and held for 0.38 min. The average helium carrier gas flow velocity was 11 cm/s. Each phase in each flask was sampled three times. 2 μL of each sample was injected three times at a 5:1 split ratio.

Table 4-9. Experimental and Calculated K_{ow} Values for Selected Chlorinated DBPs

Compound	Log Kow			
	Literature (Experimental)	Kowwin	MMPro	
			Atom Estimate	Fragment Estimate
Trihalomethanes (THMs)				
Chloroform	1.97[a]	1.52	1.61	1.95
Bromodichloromethane	2.10[a]	1.61	1.58	1.89
Bromoform	2.30[a]	1.79	1.52	2.37
Chlorodibromomethane	2.24[a]	1.70	1.55	1.83
Haloacetic Acids (HAAs)				
Monochloroacetic Acid	0.22[b,c]	0.34	0.67	1.08
Dichloroacetic Acid	0.92[b,d]	0.52	0.91	1.39
Trichloroacetic Acid	1.33[b,d]	1.44	1.50	2.10
Monobromoacetic Acid	0.41[b]	0.43	0.74	1.22
Dibromoacetic Acid		0.70	0.67	1.67
Tribromoacetic Acid		1.71	1.41	2.52
Haloacetonitriles (HANs)				
Bromoacetonitrile		0.2	1.07	-0.54
Dichloroacetonitrile (DCAN)	**-0.59***	0.29	1.25	-0.37
Trichloroacetonitrile	2.09[b]	1.2	1.84	0.34
Chloral hydrate (CH)	**1.46***	0.98	1.16	0.18

***Values determined in this experiment**

[a] Mackay et al., 1993; [b] Howard, 1997; [c] Howard, 1991; [d] Mackay et al., 1995

Kowwin and MMPro; Atom/fragment method of estimating Kow; Syracuse research Corporation

References

Brooke, D.N., A.J. Dobbs and N. Williams, 1986. Octanol: Water Partition Coefficients (P): Measurement, Estimation, and Interpretation, Particularly for Chemicals with $P>10^5$. *Ecotoxicology and Environmental Safety* **11**: 251-260.

Howard, Philip H., 1991. *Handbook of Environmental Fate and Exposure Data for Organic Chemicals*. Volume 3: Pesticides. Chelsea, Michigan, Lewis Publishers, Inc.

Howard, Philip H., 1997. *Handbook of Environmental Fate and Exposure Data for Organic Chemicals*. Volume 5: Solvents. Chelsea, Michigan, Lewis Publishers, Inc.

Mackay, Donald, Wan-Ying Shiu and Kuo-Ching Ma, 1993. *Illustrated Handbook of Physical Chemical Properties and Environmental Fate for Organic Chemicals*. Volume 3: Volatile Organic Chemicals. Boca Raton, Lewis Publishers.

Mackay, Donald, Wan-Ying Shiu and Kuo-Ching Ma, 1995. *Illustrated Handbook of Physical and Chemical Properties and Environmental Fate for Organic Chemicals*. Volume 4: Oxygen, Nitrogen, and Sulfur-Containing Compounds. New York, CRC Lewis Publishers.

U.S. EPA, 1995a. *Method 551.1 Determination of Chlorination Disinfection Byproducts, Chlorinated Solvents, and Halogenated Pesticides/Herbicides in Drinking Water by Liquid-Liquid Extraction and Gas Chromatography with Electron Capture Detection*. Cincinnati, Ohio. US Environmental Protection Agency. 1995.

U.S. EPA, 1995b. *Method 552.2 Determination of Haloacetic Acids and Dalapon in Drinking Water by Liquid-Liquid Extraction, Derivatization and Gas Chromatography with Electron Capture Detection*. Cincinnati, Ohio. US Environmental Protection Agency. 1995.

van Haelst, A.G., P.F. Heese, F.W.M. van der Wielen and H.A.J. Govers, 1994. Determination of n-Octanol/Water Partition Coefficients of Tetrachlorobenzyltoluenes Individually and in a Mixture by the Slow Stirring Method. *Chemosphere* **29**(8): 1651-1660.

Chapter 5

Potential for Exposure Misclassification and Effect on OR of Variations in DBP Concentrations between the Water Treatment Plant and Distribution System

Introduction

The difficulties in establishing accurate exposure assessment have been well-documented (see Chapter 1) in epidemiology studies investigating a possible causal association between exposure to disinfection by-products (DBPs) in treated drinking water and adverse reproductive effects. The quest for better exposure assessment is driven by the recognition that if there is indeed a causal association between DBPs and adverse reproductive effects and that the inconclusive and sometimes contradictory results obtained from epidemiology studies to date must be the result of inadequate exposure assessment. We considered the question asked at recent international workshops[1] on the subject: how good must exposure assessment data be for these epidemiology studies to be useful? This question addresses the crux of the matter with respect to exposure assessment. There are many possible levels of exposure assessment, ranging from extremely detailed, invasive, and expensive assessment, to convenience monitoring that is less expensive, but also less accurate at the individual level. Chapter 2 introduced a framework that broke down exposure assessment into three major components and grouped the reproductive epidemiology studies done to date into five categories based on their exposure assessments. This framework is a tool to evaluate whether better exposure assessment does indeed lead to stronger causal associations in these epidemiology studies.

To answer the "how good" question, we must quantitatively determine the effect of differing levels of exposure assessment on the odds ratio (OR) of an epidemiology study. A method to do just that was developed and presented in Chapter 3. However, before this method is applied to the data generated and collected for this dissertation, we must first investigate the range of detail possible in exposure assessment. Before deciding how good exposure assessment needs to be, we should look at how good it can be and how good it is now.

For the epidemiology studies in question, the optimal exposure assessment possible, in the absence of biomarkers of exposure, would enable a calculation of the administered dose[2] of the causal agent for each study subject during the critical exposure period. Let us examine what is entailed in this extreme example of comprehensive exposure assessment. First, the causal agent must be known. Daily, if not continuous, measurements of the causal agent concentrations in water must be taken at all venues where the individual study subjects are exposed to treated drinking water. Individual study subjects must keep a journal detailing all water and beverages made with water ingested, the temperature of the water, duration and frequency of showering/bathing self and others, washing laundry and dishes, swimming in treated water, and use of hot tubs or jacuzzis. In addition, concentrations of the causal agent in the air during bathing events should be measured. Aside from the logistic problems of setting up such an exposure assessment program, the execution of this program would involve great cost, result in unwieldy amounts of data, and present a huge burden on the individual study subjects. However, certain key points can be extracted from this extreme example. Accurate exposure assessment should involve measurements of all known causal agents that are temporally and spatially proximate to the individual subjects during the critical exposure period.

---

[1] An International Workshop of Exposure Assessment for Disinfection By-Products in Epidemiologic Studies, Ottawa, Ontario, May 7-10, 2000; Safety of Water Disinfection: Balancing Chemical and Microbial risks, Miami, Florida, November 1999.

[2] See definition in Chapter 1.

As discussed in Chapter 2, no epidemiology studies to date fulfil these criteria for accurate exposure assessment. The causal agent(s) responsible for the myriad reproductive effects being investigated is (are) not known. Of the studies that have measured concentrations of DBPs, only one study has measured DBP concentrations in the homes of some individual subjects. Other studies have used quarterly DBP monitoring data in the distribution system as a surrogate measure for DBP concentrations in the home. Still other studies have used DBP monitoring data from water treatment plants. While the quarterly distribution system monitoring data may be spatially near the individuals' residences, the monitoring samples are generally taken either once a quarter, or represent an average of three monthly samples. Thus, quarterly monitoring data may not account for all DBP concentration variations over the critical exposure time period. Water treatment plant monitoring may be more frequent, but it is spatially removed from the locations of individual exposure. It is evident that the exposure assessments used in the epidemiology studies to date fall short of the key points noted above as critical to good exposure assessment. Several experts have suggested that inadequate exposure assessment results in exposure misclassification (Reif et al., 1996; Nieuwenhuijsen et al., 2000). However, the main question we are considering is "How good does exposure assessment have to be?". To date it has been assumed that the inadequate exposure assessment used in epidemiology studies has attenuated the resulting ORs towards the null value, in keeping with the assumption that there is indeed a causal association and that better exposure assessment will elucidate that association. However, the potential extent of this attenuation has not been quantified. Without such quantification, it is impossible to determine the effect of the exposure misclassification resulting from inadequate exposure assessment and further, to determine how much improvement is required for exposure assessment. By employing the method introduced in Chapter 3, the effect of exposure misclassification on the OR of a hypothetical epidemiology study will be quantified.

As discussed in Chapter 1, there are three main factors contributing to exposure assessment. The first is DBP formation, including temporal and spatial variations in DBP concentrations. The second factor is the identity of the causal agent(s). The last factor is pathways of exposure to DBPs. The data generated and collected for this dissertation can shed light on the potential effects of spatial and temporal variability in DBP concentrations and also the effect of using the wrong causal agent via the methods introduced and outlined in Chapter 3.

DBP concentrations can vary substantially over distance from the water treatment plant. As described in Chapter 1, DBP concentrations vary depending on the duration of contact with residual chlorine in the distribution system, among other factors. The duration of contact, known as contact time, varies depending on how long it takes for the water to travel from the water treatment plant to the consumer's home. This depends on the distance and time the water travels, which is derived from the location of the consumer's home, water demand, pump operation, and maintenance of the system, among other variables. Exposure misclassification can occur if water treatment plant data are used as surrogate measures of in-home or distribution system DBP concentrations. DBP concentration data from City A, City B, and City C are analyzed to determine the extent of this misclassification. Daily DBP concentration data were generated through sampling programs in the distribution system, organized and executed in Cities A and B by the author. Monthly water treatment plant and distribution system monitoring data are available from the water utilities in Cities A, B, and C. A more detailed description of the sampling program and the water treatment and distribution systems in the three cities is given in Chapter 4.

These data are also used to predict the change in OR of a hypothetical epidemiology study as a result of misclassification using the methods described in Chapter 3. The data analysis in this chapter will concentrate on two individual compounds: trichloromethane (TCM) and trichloroacetic acid (TCAA), and two groups of compounds: total trihalomethanes (TTHM) and total haloacetic acids (THAA). TCM and TTHM are the most commonly measured DBPs in epidemiology studies to date. The HAAs have generated interest toxicologically as possible causal agents. TCAA is the chosen individual HAA because it is the best candidate to date for biomarkers of exposure (Froese et al., 2002, In Press; Bader et al., Submitted).

Analysis of continuous exposure data

The daily TCM and TCAA data at the water treatment plant were correlated with the TCM and TCAA data, respectively, at in-home sampling locations in both City A and City B. The monthly TCM, TTHM, TCAA, and THAA data at the water treatment plants were correlated with the TCM, TTHM, TCAA, and THAA, respectively, at monitoring sampling points in the distribution system in City B. In City A, only monthly TTHM data were available for analysis because there was no historical monitoring data available on the concentrations of other compounds in the distribution system. The City C monthly data were analyzed using both sampling points #2 and #3 as the reference locations because there is no water treatment plant in City C. Sampling point #2 was chosen as a reference location because it is the first sampling point after the first chlorination point. Sampling point #2 is at the inlet of a large, open reservoir. Average residence time in the reservoir is estimated at two to ten days during which loss of DBPs, especially volatile DBPs, may occur. Sampling point #3 is important to investigate as a sampling point as it is the most likely location for a water treatment plant, should one exist in the future. In addition, sampling location #3 is at the outflow of the open reservoir and is the second point of chlorination.

Table 5-1 shows an example of the tabulated daily data for City A. The rest of the data are presented in the Appendix. Table 5-2 presents a summary of the highest and lowest r-values for separate compounds in each city. For example, the low and high r-values for City A daily TCM are the lowest and highest of all the r-values in the tables in the Appendix for daily sampling of TCM in City A.

The sample sizes will affect the precision of the correlation coefficients and by extrapolation the precision of the estimated attenuation of the measured OR. Larger sample sizes generally have narrower confidence intervals than smaller sample sizes and are therefore more precise. OR attenuations obtained from correlation coefficients that were calculated from small sample sizes are less precise and should be interpreted with caution.

The highest and lowest r-values are given in Table 5-2 rather than the mean or median values. The focus of this table is on the best and worst values, or the range of r-values in terms of their impact on attenuating the measured OR, rather than a median measure.

Table 5-1. Correlation between daily DBP concentration data at water treatment plant #1 and distribution system locations in City A, and its effects on the odds ratio

TCM

	Correlation (r)	True Odds Ratio			
	TCM at WTP1 (Measured)	$OR_T=1.20$	$OR_T=2.00$	$OR_T=3.50$	$OR_T=7.00$
		OR_X	OR_X	OR_X	OR_X
TCM at site A01	0.67	1.13	1.59	2.31	3.66
TCM at site A02	0.77	1.15	1.70	2.62	4.46
TCM at site A03	0.32	1.06	1.24	1.49	1.85
TCM at site A04	0.33	1.06	1.26	1.52	1.92

DCAN

	Correlation (r)	True Odds Ratio			
	TCM at WTP1 (Measured)	$OR_T=1.20$	$OR_T=2.00$	$OR_T=3.50$	$OR_T=7.00$
		OR_X	OR_X	OR_X	OR_X
DCAN at site A01	0.59	1.11	1.51	2.09	3.15
DCAN at site A02	0.82	1.16	1.77	2.80	4.94
DCAN at site A03	0.46	1.09	1.37	1.77	2.42
DCAN at site A04	0.77	1.15	1.70	2.61	4.45

DCAA

	Correlation (r)	True Odds Ratio			
	TCM at WTP1 (Measured)	$OR_T=1.20$	$OR_T=2.00$	$OR_T=3.50$	$OR_T=7.00$
		OR_X	OR_X	OR_X	OR_X
DCAA at site A01	0.95	1.19	1.94	3.30	6.40
DCAA at site A02	0.91	1.18	1.88	3.13	5.90
DCAA at site A03	0.86	1.17	1.82	2.94	5.33
DCAA at site A04	0.91	1.18	1.88	3.12	5.86

TCAA

	Correlation (r)	True Odds Ratio			
	TCM at WTP1 (Measured)	$OR_T=1.20$	$OR_T=2.00$	$OR_T=3.50$	$OR_T=7.00$
		OR_X	OR_X	OR_X	OR_X
TCAA at site A01	0.67	1.13	1.59	2.30	3.65
TCAA at site A02	0.62	1.12	1.53	2.17	3.33
TCAA at site A03	0.50	1.10	1.42	1.88	2.67
TCAA at site A04	0.72	1.14	1.64	2.46	4.04

WTP1=water treatment plant #1; A01-A04 are home sampling sites

Table 5-2. Summary of low and high correlations for water treatment plant data compared to home monitoring and distribution system data

	Low r (95% CI)	**Sample size (n)**	**High r (95% CI)**	**Sample size (n)**
Daily				
City A TCM	0.32 (-0.07,0.62)	27	0.77 (0.55,0.89)	27
City A TCAA	0.48 (0.10,0.74)	25	0.72 (0.47,0.86)	27
City B TCM	-0.09 (-0.45,0.29)	28	-0.26 (-0.58,0.13)	28
City B TCAA	0.26 (0.13,0.58)	28	0.73 (0.55,0.85)	42
Monthly				
City A TTHM	0.71 (0.54,0.83)	50	0.97 (0.95,0.98)	50
City B TCM	0.94 (0.84,0.98)	17	0.97 (0.92,0.99)	17
City B TTHM	0.53 (0.18,0.76)	26	0.94 (0.87,0.97)	26
City B TCAA	0.95 (0.88,0.98)	21	0.99 (0.97,1.00)	20
City B THAA	0.96 (0.90,0.98)	22	0.99 (0.98,1.00)	22
City C TCM #2†	-0.031 (-0.49,0.44)	18	-0.73 (-0.97,0.20)	6
City C TTHM #2†	-0.016 (-0.24,0.23)	62	-0.58 (-0.74,-0.36)	51
City C TCAA #2†	0.49 (0.01,0.79)	17	0.85 (0.54,0.96)	12
City C THAA #2†	0.01 (-0.42,0.44)	21	0.81 (0.56,0.92)	19
City C TCM #3†	0.34 (-0.65,0.90)	6	0.80 (-0.28,0.99)	5
City C TTHM #3†	0.46 (0.22,0.65)	54	0.82 (0.57,0.93)	18
City C TCAA #3†	0.01 (-0.75,0.76)	7	-0.68 (-0.95,0.15)	7
City C THAA #3†	-0.01 (-0.71,0.70)	8	0.59 (-0.20,0.91)	8

†City C data were analyzed using both sampling points #2 and #3 as reference points for comparison

There are several possible explanations for the large spread in sample sizes in Table 5-2. The monthly monitoring data were taken over a period of years, as discussed in Chapter 4. However, sampling for some DBP compounds started earlier than sampling for others, resulting in more data points for some compounds than for others. For example, in the three cities whose data are used here, sampling for TTHMs started before sampling for individual THMs, individual HAAs, or total HAAs. As a result, there are more TTHM data points available for analysis. In addition, distribution sampling points may change over time, as in the case of City C. Therefore, some sampling points were added later than the original points and have fewer data points for analysis. In the daily data, missing data points from the home are a result of the volunteers' forgetting to take the sample on a particular day. This only occurred in City B daily data. In City A, a few water treatment plant samples could not be taken because one of the plants was shut down for maintenance. Less weight can be put on those correlations obtained from small sample sizes. It should be noted that in these data sets most correlations from small sample sizes are not statistically significant.

Some general trends are seen from Table 5-2. HAA data from the water treatment plants predict HAA concentrations in the home or distribution system better than THM data from the water treatment plants predict THM concentrations in the home or distribution system. The exception is City C #3 data. For the City C #3 monthly data, both high and low TCM and TTHM r-values are higher than the high and low TCAA and THAA r-values. A second trend seen in Table 5-2 is that individual compounds have better correlations than total groups of compounds. For example, r-values for City B monthly TCMs are higher than those of the City B monthly TTHMs.

There are some negative correlations that must be discussed. Both low and high r-values for the City B daily TCM data are negative, but non-significant. City C #2 and #3 monthly data show several negative correlations. Of particular interest is the statistically significant high r-value for City C #2 TTHM (r=-0.58; 95% CI -0.74, -0.36; n=51). A negative correlation suggests that the variation behaviour of DBP concentrations at the distribution system point in question is opposite to the variation behaviour of DBP concentrations at the reference location (sampling point #2). This phenomenon likely occurs because of the large, open reservoir between sampling points #2 and #3. The sampling location at which this r-value occurs in the City C #2 analysis is at sampling point #3 (the exit point of the open reservoir). This suggests that there is indeed some loss of DBPs in the open reservoir, an observation that is supported by the summary statistics between City C sampling locations #2 and #3, shown in Chapter 4.

Interestingly, negative correlations for the City C #3 analysis only occur for the HAAs, whereas most of the negative correlations for the City C #2 analysis occur for the THMs, most of which are not significant due to small sample sizes. This may be explained by reaction kinetics and differences in the behaviour of HAAs and THMs in the City C distribution system. Chlorine is added at multiple points in the City C distribution system. The first two additions are at sampling points #1 (at the intake before travel through an aqueduct) and #3. Between sampling points #1 and #2 substantial amounts of DBPs are formed because of the high organic matter concentration in the source water and the long contact time through the aqueduct. Roughly half of the DBPs are lost in the open reservoir between sampling points #2 and #3 (Table 4-6). However, the DBP concentrations increase after sampling point #3 because of the addition of chlorine at this point. As discussed in Chapter 1, high chlorine doses favour the formation of HAAs (Krasner, 1999). In addition, HAA formation kinetics are faster than THM formation kinetics (Singer, 1999). These two factors suggest that the HAA concentrations will rise to pre-sampling point #3 levels more quickly than will the THM concentrations. This is observed in the summary statistics in Chapter 4. In general, the HAA concentrations return to sampling point #2 levels by sampling point #7. From sampling point #7 on, the HAA concentrations are fairly stable at sampling point #2 levels. THM concentrations, however, do not exhibit the same rapid return to sampling point #2 levels. In fact, THM concentrations reach sampling point #2 levels only by sampling point #10, after which they continue to rise. The rapid increase and leveling out of HAA concentrations soon after the second chlorination suggests that they reach a plateau shortly after chlorination. This same concentration plateau was reached at sampling point #2 after the long residence time in the aqueduct. Hence, HAA concentrations at sampling point #2 may be indicative of HAA concentrations in the rest of the distribution system once the HAA concentration peak has been reached after additional chlorination. However, the THMs behave in a different manner. Slower formation kinetics seem to emphasize the effect of the substantial decrease in THM concentration in the open reservoir. As a result, sampling point #3, which is after the reservoir and the next point of chlorination, acts as a second starting point from which THM concentrations can increase. The substantial loss of THMs in the open reservoir makes sampling point #2 a less desirable point from which to predict THM concentrations in the distribution system. It should also be noted that as the THM concentrations increase in the distribution system, the negative correlations with sampling point #2 concentrations are replaced by statistically significant positive correlations. However, the correlations are never quite as large as those for sampling point #3.

The data in Table 5-2 and in the appendix suggest that monthly water treatment plant data correlate better with monthly distribution system data than daily water treatment plant data do with daily distribution system data. In other words, if water treatment plant data are used to predict distribution system data, the data suggest that monthly rather than daily data should be used. This result is counterintuitive to what would be expected for good exposure assessment at the individual level. While monthly data may give better correlation coefficients, daily data are still likely to assess exposure better at the individual level. In addition, it should be noted that these results apply for only the water systems investigated in this study. These results have not been shown to be generalizable over many water systems and only will by virtue of repetition of this method over multiple systems. The greater variability shown by the daily data over the monthly data could result from differences in the sampling methods. Monthly samples were taken by trained operators at the water treatment plant, likely resulting limited introduction of error at the sampling stage. The daily samples were taken by volunteers, some of whom were laboratory-trained, but not all. There was likely more introduction of error and variability between samples introduced at the sampling stage in the daily sampling than in the monthly sampling.

There is evidence that chloraminated systems have the advantage over chlorinated systems in correlating distribution system concentrations with water treatment plant concentrations. City A daily data (chloraminated system) have higher r-values than City B daily data (chlorinated system). Additionally, City A daily data have more r-values that are statistically significant at the 95% level. City A monthly TTHM correlations are generally higher than City B monthly TTHM data. As mentioned in Chapter 4, the predominant difference between Cities A and B is the type of disinfection residual (chloramine and chlorine, respectively).

From the correlations between water treatment plant and distribution system samples we have inferred the potential for exposure misclassification if water treatment plant samples are used as surrogate measures for distribution system samples. Ultimately, we need to know the effect of this misclassification on the OR of a study, which we can determine using the methods outlined in Chapter 3. For the purposes of the analysis method, we will assume that the distribution system and home data represent the "true" data sets and the water treatment plant data represent the "measured" data sets. Measured ORs were calculated for the water treatment plant data for four "true" ORs: 1.20, 2.00, 3.50, and 7.00. For the purposes of this exercise, the "true" ORs are considered the result that would occur if the "true" distribution system/home data were used as the exposure measure. These data are in the appendix.

As a brief reminder, the method introduced in Chapter 3 mathematically linked OR attenuation and the correlation coefficient between "true" and "measured" data sets using the attenuation equation:

$$\beta_X = \rho^2_{TX}\beta_T$$

Where β_X and β_T are the logistic regression coefficients for the measured and true data, respectively, and ρ is the correlation coefficient between the true and measured data.

The expression for the odds ratio for the true and measured data are:

$$OR_T = \exp(\beta_T s \sigma_T)$$
$$OR_X = \exp(\beta_X s \sigma_X)$$

Where σ_T and σ_X are the standard deviations in T and X, respectively, and s is the number of standard deviations (see Chapter 3).

Substituting the attenuation equation into the logistic regression equation gives the measured OR as a function of the true OR and the correlation coefficient:

$$OR_X = OR_T \exp(\rho_{TX})$$

This last equation is used to predict the OR attenuation as a result of exposure misclassification as evidenced by r differing from 1.00.

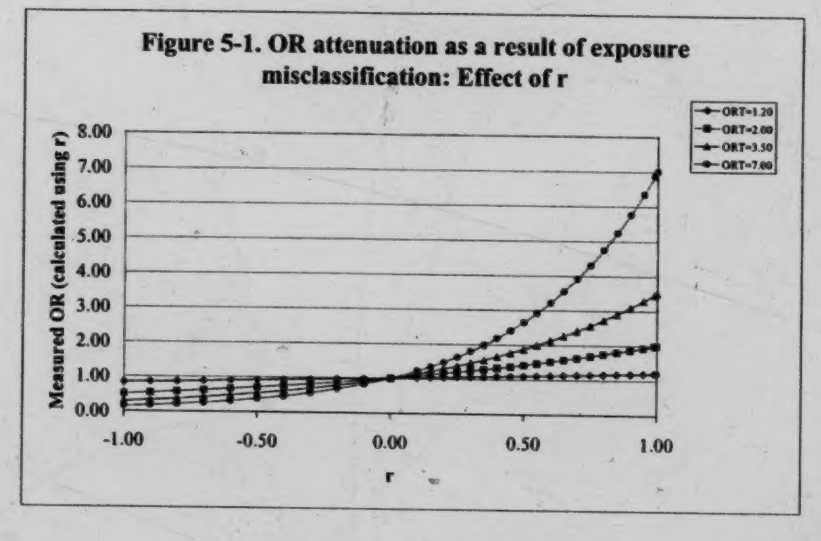


Figure 5-1. OR attenuation as a result of exposure misclassification: Effect of r

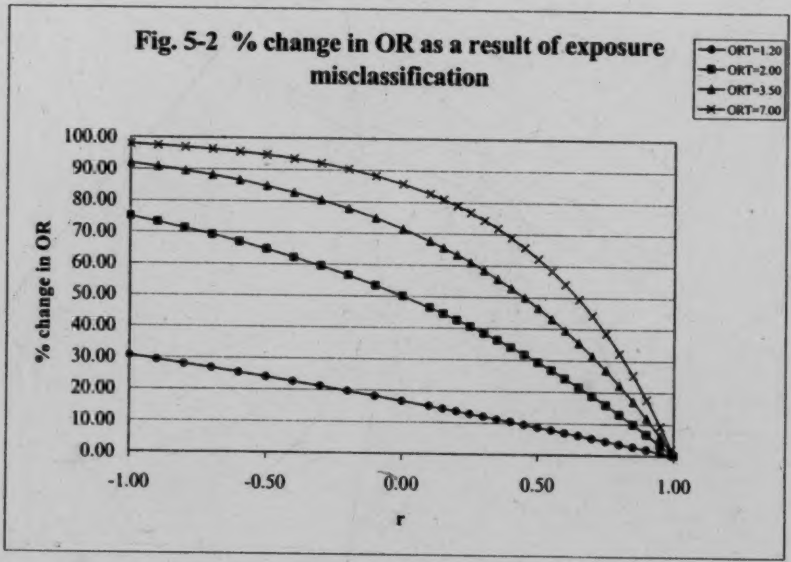


Fig. 5-2 % change in OR as a result of exposure misclassification

Figure 5-1 shows OR attenuation as a result of exposure misclassification. Figure 5-2 shows the percent change in OR (calculated as: $\{[OR_T\text{-}OR_X]/OR_T\}$ X 100) as a result of exposure misclassification. These graphs are derived from the OR equations presented in Chapter 3 applied to "true" ORs of 1.20, 2.00, 3.50, and 7.00.

The graphs indicate that even slight deviations of the r-values from 1.00 can have a substantial effect on the OR. In addition, the extreme attenuation associated with low r-values is evident. The City C monthly #2 r-value for TTHM (r=-0.58; 95%CI -0.74, -0.36; n=51) is an example of the extreme attenuation associated with negative r-values. For example, at OR_T=1.20, the measured OR is 0.90. At OR_T=2.00, the measured OR is 0.67. At OR_T=3.50, the measured OR is 0.48. Finally, at OR_T=7.00, the measured OR is 0.32. The "true" ORs are greater than one (OR_T = 1.20, 2.00, 3.50, or 7.00), suggesting in this hypothetical study that there is a causal association between the agent and the adverse effect. An OR less than one in this case suggests a *protective* effect from exposure to the agent. While the TTHM concentrations at the reference point are associated with TTHM concentrations at the distribution system point, data from each location would give opposite results pertaining to a causal association should they be used in an epidemiology study. This is an extreme example of the effect exposure misclassification can have on the results of an epidemiology study. Examples of less substantial OR attenuation associated with higher r-values are found in the data tables in the appendix.

We have discussed the potential for exposure misclassification and its effects on OR attenuation. Now we can attempt an answer to the question "How good does the exposure assessment have to be?". First, assume that we want the measured data to give a fairly close estimate of the true OR, say, within 10%. Given this parameter, what level of correlation would be needed between the measured data and the "true" data? This measure depends on the "true" OR. For OR_T of 7.00, r-values of 0.95 or greater are required for the measured OR to be within 10% of the OR_T. For OR_T of 3.50, r-values of 0.92 or greater are required. For OR_T of 2.00, r-values of 0.85 or greater are required. For OR_T = 1.2, r-values of only 0.42 or greater are required. As the "true" OR increases, the measured data must be correlated more closely to the "true" data to produce an OR within 10% of the "true" OR. A summary of the number of measured data sets that result in the required r-values for each city and sampling frequency is presented in Table 5-3 and Figures 5-3 and 5-4.

Table 5-3. Number of data sets with r-values resulting in measured ORs (OR_X) within 10% of the true ORs (OR_T)

Location	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00	
	# of data sets with $r \geq 0.42$	# of data sets with $r \geq 0.85$	# of data sets with $r \geq 0.92$	# of data sets with $r \geq 0.95$	Total # data sets
Daily					
City A volatiles: Total (WTP1; WTP2)	13 (6;7)	0	0	0	16
City A HAAs: Total (WTP1;WTP2)	16 (8;8)	7 (4;3)	1	1	16
City B THMs	0	0	0	0	8
City B HAAs	3	0	0	0	8
Monthly					
City A THMs (WTP1;WTP2)	24 (12;12)	10 (3;7)	7 (0;7)	6 (0;6)	24
City B THMs	12	6	6	4	12
City B HAAs	12	12	12	12	12
City C THMs #2[†]	4	0	0	0	39
City C HAAs #2[†]	30	1	0	0	39
City C THMs #3[†]	32	0	0	0	36
City C HAAs #3[†]	5	0	0	0	36

[†] City C data were analyzed using both sampling points #2 and #3 as reference points

WTP1=Water treatment plant #1; WTP2 = Water treatment plant #2

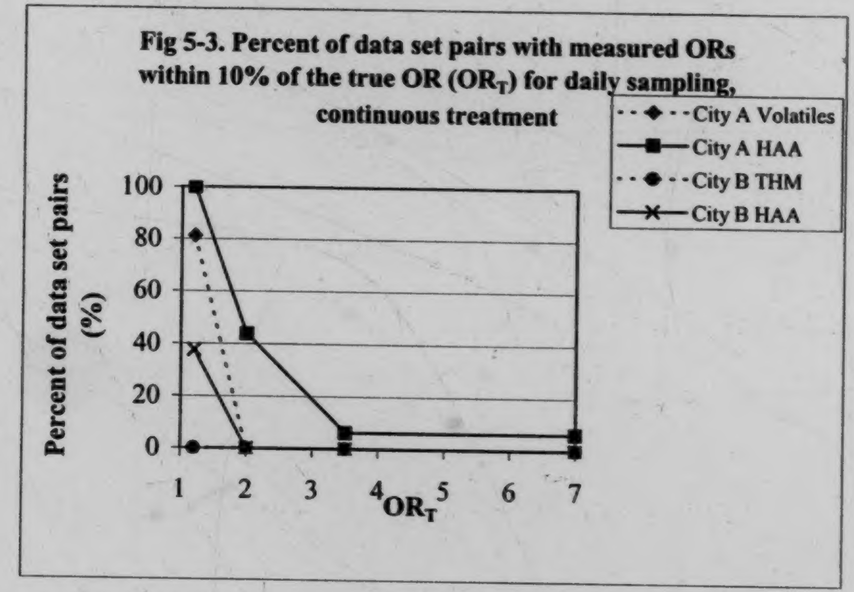


Fig 5-3. Percent of data set pairs with measured ORs within 10% of the true OR (OR_T) for daily sampling, continuous treatment

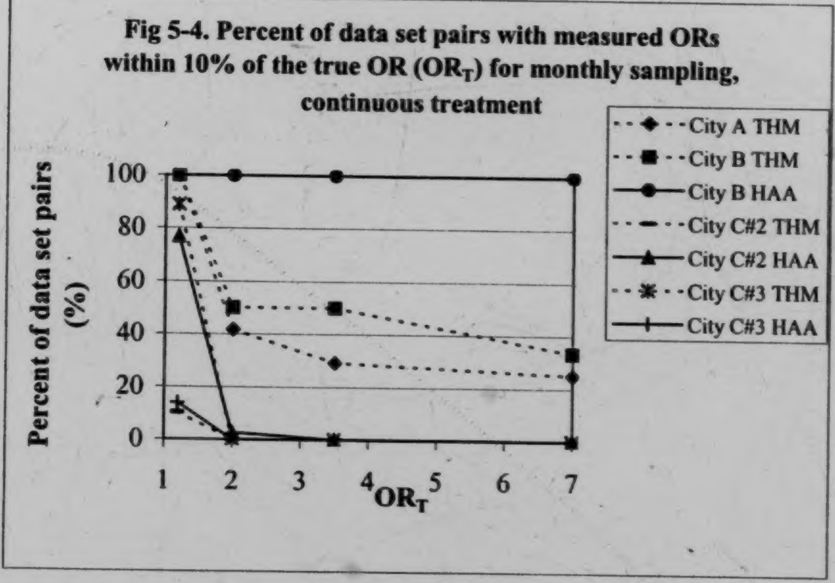


Fig 5-4. Percent of data set pairs with measured ORs within 10% of the true OR (OR_T) for monthly sampling, continuous treatment

The first observation about these data is that, generally, HAA data sets tend to produce more r-values within the critical range than do the non-HAA data sets. This supports the observation made earlier that HAAs are better predictors than THMs if water treatment plant data are used as surrogate measures of DBP concentrations in the distribution system. This particularly holds if the true OR is expected to be quite large. The exception to this observation is the City C #3 data where the THMs produced more r-values within the critical range than the HAAs. This issue was discussed previously and explained by the reference point locations with respect to an open reservoir and DBP formation kinetics.

In addition, the observation that water treatment plant data can better be used as surrogate measures for distribution system data in chloraminated systems than in chlorinated systems was clearly demonstrated for daily sampling. Neither system produced qualifying r-values at the three highest OR_Ts for the THM data sets. However, for $OR_T=1.20$, 13 of 16 (81%) possible City A THM data sets exhibited the requisite r-values, whereas no City B THM data sets qualified. Three of the eight (38%) possible City B HAA data sets resulted in qualifying r-values at $OR_T=1.20$. All of the 16 (100%) possible City A HAA data sets were within the requisite range for $OR_T=1.20$, with 7 (44%) qualifying at $OR_T=2.00$, and one (6%) data set qualifying in each of the two highest OR_T values. The monthly sampling THM data show that both the chlorinated and chloraminated systems exhibit qualifying r-values at all OR_Ts. However, the superiority of the chloraminated system over the chlorinated system was not clearly demonstrated in the monthly sampling.

In analyzing these data, we may find it useful to reword our original question to: "Can we use water treatment plant DBP concentration data as a surrogate measure of DBP concentrations in the distribution system and still get a reasonable estimate of the 'true' OR?" The answer depends on the agent measured, the frequency of sampling, the expected "true" OR, and the type of disinfection residual used. HAAs are better agents to measure than THMs in this situation, regardless of other factors. This is particularly true at low "true" ORs. However, THM data were generally adequate in the chloraminated system for both daily and monthly sampling frequencies. THM data at the water treatment plant (or reference sampling point) were useful neither in the City C chlorinated system for monthly sampling nor in the City B chlorinated system for daily sampling.

In general, daily sampling provides fewer r-values within the requisite range than monthly sampling. In other words, daily DBP concentrations at the water treatment plant do not provide as good surrogate measures for daily DBP concentrations in the distribution system as monthly DBP concentrations at the water treatment plant do for monthly DBP concentrations in the distribution system. This particularly applies to chlorinated systems and when measuring THMs for higher expected true ORs. The exception to this observation is the daily measurement of HAAs in a chloraminated system, particularly with a low expected "true" OR. In this case the water treatment plant and distribution system data sets correlated to produce r-values within the required range to result in a larger proportion of ORs within 10% of the "true" ORs. However, if exposure assessment is based on monthly samples, then it may not be worth the additional cost and effort to measure DBP concentrations in the distribution system. The monthly water treatment plant DBP concentration data seem to be fairly good surrogate measures of monthly distribution system DBP concentration data in both the chlorinated and chloraminated systems with water treatment plants and for both THMs and HAAs.

Analysis of categorical exposure data

Many epidemiology studies group exposure data into categories rather than using continuous data. This is often done to account for imprecisions in the continuous data from sampling and analysis. However, the differences in exposure misclassification between categorical and continuous treatment of exposure data may result in different levels of OR attenuation.

All the exposure data in these analyses were divided into quartiles. The top quartile was designated as "exposed" while the lower three quartiles were designated as "not exposed". As with the continuous treatment, "true" and "measured" data were required for the categorical treatment methods. The water treatment plant data sets were designated as the "measured" data, and the distribution system data sets were designated as the "true" data, as in the continuous treatment. Sensitivity and specificity were calculated for each pair of measured/true data sets. Sensitivity and specificity are epidemiological measures of misclassification; in this case the misclassification is due to exposure assessment. Sensitivity is the proportion of study subjects who are truly exposed and are classified as exposed. Specificity is the proportion of study subjects who are truly not exposed and are classified as not exposed. Examples of the calculation of sensitivity and specificity are provided in Chapter 1. In Chapter 3 we introduced equations that link exposure misclassification, as measured by sensitivity and specificity, to OR attenuation. As a reminder, these equations are presented below.

The main equation is:

$$OR_X = [p_D(1-p_N)]/[p_N(1-p_D)]$$

Where:

$$p_D = sens_D P_D + (1-spec_D)(1-P_D)$$
$$p_N = sens_N P_N + (1-spec_N)(1-P_N)$$

$sens_N$ and $sens_D$ are the sensitivity for the non-diseased and diseased groups, respectively, and $spec_N$ and $spec_D$ are the specificity for the non-diseased and diseased groups, respectively. Since we are interested only in non-differential exposure misclassification, $sens_D = sens_N$ and $spec_D = spec_N$.

P_N is the proportion of exposed subjects who do not have the disease and is set arbitrarily in the calculations here to be 0.2. P_D, another value required for the development of the tables, is from the equation:

$$P_D = P_N OR_T/[1+P_N(OR_T-1)]$$

An in-depth explanation of these equations is provided in Chapter 3.

Table 5-4 is an example of the tabulated sensitivity, specificity, and attenuated OR data from City A daily sampling. The rest of the data are presented in the Appendix for Chapter 5.

Table 5-4. Sensitivity and specificity between daily DBP concentration data at water treatment plant #1 and distribution system locations in City A, and their effects on the odds ratio

TCM

			True Odds Ratio			
	TCM at WTP1 (Measured)		$OR_T=1.20$	$OR_T=2.00$	$OR_T=3.50$	$OR_T=7.00$
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
TCM at site A01	0.60	0.86	1.08	1.38	1.83	2.55
TCM at site A02	0.33	0.79	1.02	1.10	1.19	1.33
TCM at site A03	0.09	0.69	0.97	0.86	0.72	0.56
TCM at site A04	0.67	0.90	1.11	1.51	2.14	3.20

DCAN

			True Odds Ratio			
	DCAN at WTP1 (Measured)		$OR_T=1.20$	$OR_T=2.00$	$OR_T=3.50$	$OR_T=7.00$
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
DCAN at site A01	0.60	0.86	1.08	1.38	1.83	2.55
DCAN at site A02	0.83	0.95	1.15	1.74	2.75	4.70
DCAN at site A03	0.29	0.80	1.02	1.07	1.14	1.23
DCAN at site A04	0.67	1.00	1.18	1.86	2.94	4.79

DCAA

			True Odds Ratio			
	DCAA at WTP1 (Measured)		$OR_T=1.20$	$OR_T=2.00$	$OR_T=3.50$	$OR_T=7.00$
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
DCAA at site A01	1.00	1.00	1.20	2.00	3.50	7.00
DCAA at site A02	1.00	0.91	1.15	1.73	2.83	5.40
DCAA at site A03	0.86	0.95	1.15	1.75	2.78	4.82
DCAA at site A04	1.00	0.87	1.13	1.66	2.64	4.94

TCAA

			True Odds Ratio			
	TCAA at WTP1 (Measured)		$OR_T=1.20$	$OR_T=2.00$	$OR_T=3.50$	$OR_T=7.00$
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
TCAA at site A01	0.75	0.80	1.08	1.38	1.87	2.72
TCAA at site A02	1.00	0.79	1.11	1.54	2.36	4.26
TCAA at site A03	0.50	0.60	1.01	1.06	1.11	1.19
TCAA at site A04	0.91	0.88	1.12	1.60	2.46	4.23

The same general trends are seen in the categorical data treatment as in the continuous data treatment. In fact, trends are more pronounced in the categorical treatment. As in the continuous treatment, HAA measurements at the water treatment plants are better surrogate measures of HAAs in the distribution system than THM water treatment measurements are of THM distribution system measurements. DBP concentrations at the water treatment plants have a better correlation with distribution system DBP concentrations in chloraminated systems than in chlorinated systems. However, in chlorinated systems, monthly water treatment plant data predict monthly distribution system data better than daily water treatment data predict daily distribution system data. The most noticeable discrepancy between the categorical and continuous treatments is the fact that the measured ORs in the categorical treatment are almost always lower than the corresponding measured ORs in the continuous treatment. This issue will be discussed in more detail in Chapter 8.

A few notes need to be addressed regarding some of the data. For multiple points in the City C #2 and #3 data, the measured OR could not be calculated because the sensitivity could not be calculated. Remember that sensitivity is the proportion of subjects who are truly exposed that are classified as exposed. In mathematical terms this can be written as:

Sensitivity = a/a+c

Where:
a = the number of subjects truly exposed and classified as exposed
c = the number of subjects truly exposed and classified as not exposed
a+c = the total number of subjects truly exposed

The only case in which sensitivity can not be calculated is if no subjects are truly exposed. In this case, the denominator of the equation will be zero. This situation is likely to occur only when the sample size (n) is very small. As mentioned in a previous discussion, sampling locations and the compounds monitored in City C have changed over time. In the data set pairing process to calculate sensitivity and specificity, those points with missing data in either data set of the data set pair were discarded from the analysis. For some data set pairs, this resulted in very low sample sizes, and subsequent calculation difficulties.

As in the continuous analysis, there are several instances where the OR attenuation results in measured ORs less than one ($OR_X<1.00$). OR attenuations to below one occur when sensitivity is at or near zero, or both sensitivity and specificity are low. Low values for the sensitivity and specificity indicate that there is substantial exposure misclassification. In the cases where the OR is attenuated below 1.00, the misclassification is so extreme that the resulting OR suggests the opposite result from the true OR. For example, if a positive true OR suggests a causal effect between an exposure and an outcome, then a negative measured OR suggests a protective effect from the exposure. The most extreme case of misclassification, in which both sensitivity and specificity are zero, does not occur in these data. There are several examples of OR attenuation below 1.00. City C monthly data have the most examples of measured ORs below one. Again, there is a difference between the DBP species for each reference point. For reference point #2, the THMs exhibit the most misclassification as measured by low sensitivities and specificities and attenuated ORs. For reference point #3, the HAAs exhibit the most misclassification. This is similar to the observation made in the continuous analysis which suggested that more exposure misclassification occurs with THMs when distribution system values are compared to reference point #2, whereas more misclassification occurs with HAAs when distribution system values are compared to reference point #3.

The purpose of this analysis is to answer the question: "How good does exposure assessment have to be?". In the continuous data treatment, we determined the correlation required between the "true" and "measured" data sets to result in an OR_X within 10% of the OR_T. However, with the categorical analysis, we have two independent variables with complex interactions affecting the OR_X. As a result, it is difficult to produce a simple table to describe the sensitivity and specificity requirements to produce an OR_X within 10% of the OR_T. Therefore, Table 5-5 and Figures 5-5 and 5-6 show the number of data sets within each city and sampling frequency that result in an OR_X within the required range, based only on the final OR_X and not on the sensitivity and specificity.

Table 5-5. Number of measured ORs within 10% of the true OR

Location	$OR_T=1.20$	$OR_T=2.00$	$OR_T=3.50$	$OR_T=7.00$	
	# of data sets with $OR_X \geq 1.08$	# of data sets with $OR_X \geq 1.80$	# of data sets with $OR_X \geq 3.15$	# of data sets with $OR_X \geq 6.30$	Total # data sets
Daily					
City A volatiles: Total (WTP1; WTP2)	6 (5;1)	0	0	0	16
City A HAAs: Total (WTP1;WTP2)	12 (6;6)	2 (1;1)	2 (1;1)	1 (1;0)	16
City B THMs	0	0	0	0	8
City B HAAs	1	0	0	0	8
Monthly					
City A THMs (WTP1;WTP2)	23 (12;11)	4 (1;3)	2 (0;2)	0 (0;0)	24
City B THMs	12	2	0	0	12
City B HAAs	12	9	9	0	12
City C THMs #2[†]	0	0	0	0	39
City C HAAs #2[†]	15	2	0	0	39
City C THMs #3[†]	16	9	7	6	36
City C HAAs #3[†]	6	0	0	0	36

[†] City C data were analyzed using both sampling points #2 and #3 as reference points for comparison; WTP1=Water treatment plant #1 water treatment plant; WTP2 = Water treatment plant #2 water treatment plant

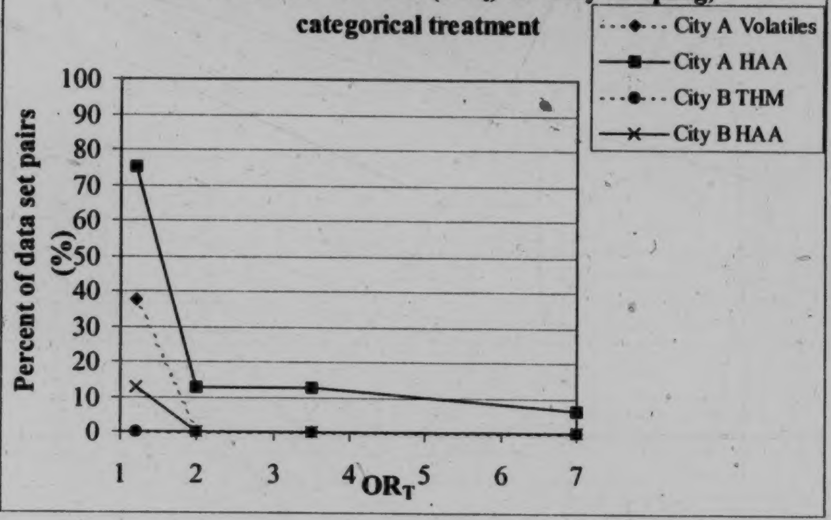


Fig 5-5. Percent of data set pairs with measured ORs within 10% of the true OR (OR_T) for daily sampling, categorical treatment

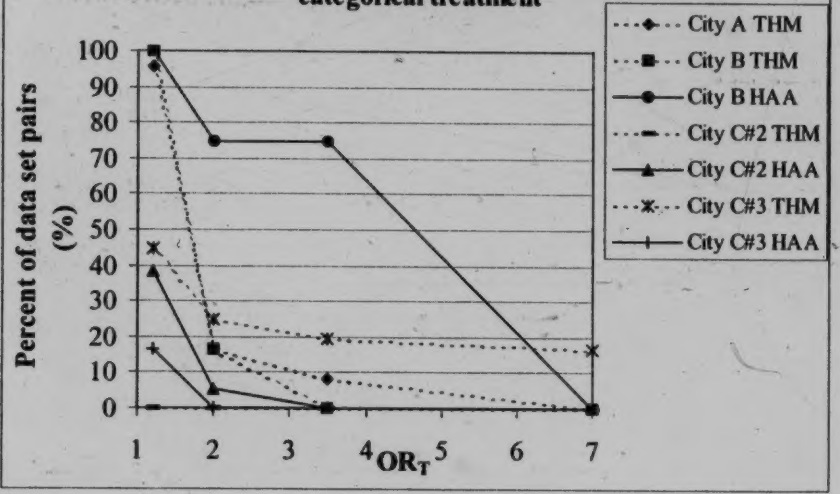


Fig 5-6. Percent of data set pairs with measured ORs within 10% of the true OR (OR_T) for monthly sampling, categorical treatment

Summary and Conclusions

In this chapter we investigated the question "How good does exposure assessment have to be?" in relation to the variations in DBP concentrations between the water treatment plant and the distribution system. We looked at measures of misclassification and were able to draw some conclusions about factors that affect DBP concentration variation and exposure misclassification. We treated the data both categorically and continuously, and found similar trends in both treatments, although categorical measured ORs tended to be lower than continuous measured ORs.

Several general observations from these analyses can reduce exposure misclassification. Several factors, including the compounds measured, the type of disinfection, and the frequency of sampling, clearly contributed to misclassification and as such imply methods of improving exposure assessment.

The distribution system DBP concentrations correlated better to water treatment plant DBP concentrations in a chloraminated system than a chlorinated system. In addition, the chloraminated system produced more measured ORs within 10% of the "true" OR than did the chlorinated data sets for both the categorical and continuous analyses.

Correlations were higher between distribution system HAAs and water treatment plant HAAs than between distribution system THMs or DCAN and water treatment plant THMs or DCAN. Therefore, if water treatment plant data are used to predict distribution system concentrations, HAAs should be the compounds measured. In terms of minimizing the effect of exposure misclassification on the OR of an epidemiology study due to DBP concentration variations between the distribution system and the water treatment plant, HAAs performed better than THMs or DCAN, particularly at higher "true" ORs. This observation applies to both the continuous and the categorical treatment of the data. However, the effect was more pronounced in the categorical treatment of the data. The HAA data produced more measured ORs within 10% of the "true" ORs than did the THM or DCAN data. The exception is City C #3, which is a result of certain distribution characteristics unique to City C.

For the most part, the water treatment plant data predicted the and distribution system data better with monthly than with daily sampling. In addition, the monthly data produced more measured ORs within 10% of the "true" OR than did the daily data. This observation applies to both the continuous and categorical data treatments. It should be noted that these results apply for only the water systems investigated in this study and are not necessarily generalizable to other water systems. The greater variability shown by the daily data over the monthly data could result from differences in the sampling methods.

Ideally, detailed and specific exposure assessments for each study subject would form the basis of the exposure measure in epidemiology studies looking at a possible causal association between exposure to DBPs in drinking water and adverse reproductive effects. However, resource and feasibility constraints ensure that this will not be achieved for a full-scale field study. Therefore, from the general observations taken from these analyses of the data available for this study, we can make a general statement aimed towards minimizing exposure misclassification for these studies. If water treatment plant DBP concentrations are used as surrogate measures for distribution system DBP concentrations, then the least misclassification and least resulting OR attenuation is likely to occur if *monthly HAA data* from a *chloraminated system* are used. As discussed in this chapter, there are exceptions to this general statement. However, if non-HAA compounds are measured in systems using chlorination, the data support current thinking that exposure misclassification due to spatial variations in DBP concentrations is best avoided by sample collection as close to the individual study subjects as possible.

References

Bader, E., K. F. Froese, S.E. Hrudey Submitted. Determination of Elimination Half-Life for TCAA in a Controlled Human Exposure Study. .

Froese, K.F., M.I. Sinclair, S.E. Hrudey 2002, In Press. Trichloroacetic Acid as a Biomarker of Human Exposure to DBPs in Drinking Water - Adelaide Trial. *Environmental Health Perspectives* **110**.

Krasner, S. 1999. Chemistry of Disinfection By-Product Formation. in *Formation and control of Disinfection By-Products in Drinking Water*. P. C. Singer. Denver, Colorado, American Water Works Association: 424.

Nieuwenhuijsen, M.J., M.B. Toledano, N.E. Eaton, J. Fawell, P. Elliott 2000. Chlorination Disinfection By-Products in Water and Their Association with Adverse Reproductive Outcomes: A Review. *Occupational and Environmental Medicine* **57**: 73-85.

Reif, J.S., M.C. Hatch, M. Bracken, L.B. Holmes, B.A. Schwetz, P.C. Singer et al., 1996. Reproductive and Developmental Effects of Disinfection By-Products in Drinking Water. *Environmental Health Perspectives* **104**(10): 1056-1061.

Singer, P.C. 1999. *Variability and Assessment of Disinfection By-Product Concentrations in Water Distribution Systems*. The Safety of Water Disinfection: Balancing Chemical and Microbial Risks, Miami Beach, Florida, International Life Sciences Institute.

Chapter 6

Potential for Exposure Misclassification and Subsequent OR Attenuation from Measuring TCM or TTHM rather than More Toxicologically Plausible Causal Agents

Introduction

Chapter 1 provided an introduction to epidemiology studies investigating a possible causal association between disinfection by-products (DBPs) in treated drinking water and adverse reproductive and developmental effects. The challenges associated with exposure assessment in these studies were discussed. Three main factors affecting exposure assessment were presented: DBP formation, identity of causal agent, and route of uptake. In Chapter 5 we presented an extreme example of exposure assessment to DBPs in treated drinking water. From this example, we extracted several key points crucial to good exposure assessment for these studies, including measurement of the correct causal agent and measurement of DBP concentrations temporally and spatially proximate to individual study subjects during the critical exposure periods.

In Chapter 5, we investigated the effect of spatial and temporal variability of DBP concentrations on the measured odds ratio (OR) of a hypothetical epidemiology study. In doing so, we shed light on the answer to the question "How good does exposure assessment have to be?" in relation to DBP concentration variations in the distribution system. In this chapter we investigate the effect on exposure misclassification produced by measuring agents that may not be the causal agents. Those epidemiology studies to date that have measured DBP concentrations have measured THMs almost exclusively. As discussed in Chapter 1, however, toxicology studies suggest that other DBP species, in particular the HAAs, have more potential to be causal agents for adverse reproductive effects. In this chapter we will determine the potential for exposure misclassification from measuring THM rather than other DBP species. For example, if TCAA is a causal agent, how much exposure misclassification will occur if TCM concentrations are used as the exposure measures instead of TCAA concentrations? In addition, we will quantify the OR attenuation in a hypothetical study as a result of any exposure misclassification. In essence, we are endeavoring to answer the question "how good does exposure assessment have to be?" from the angle of measuring correct causal agents or compounds that are not the causal agents. Since there is uncertainty regarding the identity of the true causal agents(s), it is useful to discover the potential ramifications of measuring DBPs other than THMs. The statistical methods introduced in Chapter 3 and employed in Chapter 5 will be used to determine the potential for exposure misclassification and the OR attenuation as a result of that misclassification. The raw data used in this chapter, including the daily data from Cities A and B and the monthly monitoring data from Cities A, B, and C are the same as those used in Chapter 5. In addition, correlations from published studies will be incorporated into the continuous exposure analysis. As in Chapter 5, the data will be treated as both continuous and categorical.

Analysis of continuous exposure data

The daily TCM data were correlated against other DBP species at each water treatment plant and each home sampling point in Cities A and B. The monthly TCM and TTHM monitoring data were correlated against other DBP species data at each water treatment plant and at each distribution system sampling point in Cities B and C. In City A, only the water treatment plant data were used because monthly monitoring data for other DBPs than TCM and TTHM were not available in the distribution system. Table 6-1 is an example of the correlations and attenuated ORs for the City A daily data. The rest of the data are presented in the Appendix. Table 6-2 is a summary of the high and low correlations for all the data sets. The highest and lowest values of r from the tables in the Appendix were used in Table 6-2. As explained in Chapter 3, the statistics demand that in each correlated data set pair, one data set must be designated "measured, and one must be designated "true". TCM and TTHM were chosen as the "measured" data sets because they are most commonly used DBP species in epidemiology studies, although there have been little toxicological data to support their potential as causal agents. The actual causal agent(s) is (are) not known, therefore, all other DBP species besides TCM and TTHM for which data were available were considered potential causal agents and designated as the "true" data sets. The "true" data sets for the purposes of our analysis included BDCM, DCAA, TCAA, THAA, and DCAN. Data on all these compounds were not available at all locations. A more detailed description of the sampling program, data collection and water treatment systems in the three cities is found in Chapter 4. All correlations and calculations between TCM and TTHM were near or at r=1.00 and resulted in little or no OR attenuation. Therefore, these results are not included in the tables or the discussion. High correlations between TCM and TTHM is expected because TCM makes up the bulk of TTHM concentrations.

The sample sizes will affect the precision of the correlation coefficients and by extrapolation the precision of the estimated attenuation of the measured OR. Larger sample sizes generally have narrower confidence intervals than smaller sample sizes and are therefore more precise. OR attenuations obtained from correlation coefficients that were calculated from small sample sizes are less precise and should be interpreted with caution.

The highest and lowest r-values are given in Table 6-2 rather than the mean or median values. The focus of this table is on the best and worst values, or the range of r-values in terms of their impact on attenuating the measured OR, rather than a median measure.

Table 6-1. Correlation between TCM and other DBP species at water treatment plants 1 and 2 and distribution system locations in City A, and its effects on the odds ratio

City A Daily

	Correlation (r)	True Odds Ratio			
	TCM at WTP1 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
DCAN at WTP1	0.92	1.18	1.89	3.17	6.01
DCAA at WTP1	0.79	1.16	1.73	2.70	4.69
TCAA at WTP1	0.89	1.18	1.85	3.04	5.62

	Correlation (r)	True Odds Ratio			
	TCM at WTP2 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
DCAN at WTP2	0.86	1.17	1.81	2.93	5.31
DCAA at WTP2	0.55	1.11	1.47	2.00	2.94
TCAA at WTP2	0.66	1.13	1.58	2.30	3.64

	Correlation (r)	True Odds Ratio			
	TCM at A01 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
DCAN at site A01	0.84	1.17	1.79	2.86	5.11
DCAA at site A01	0.65	1.13	1.57	2.26	3.56
TCAA at site A01	0.83	1.16	1.77	2.82	5.00

	Correlation (r)	True Odds Ratio			
	TCM at A02 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
DCAN at site A02	0.84	1.16	1.79	2.85	5.10
DCAA at site A02	0.62	1.12	1.54	2.18	3.35
TCAA at site A02	0.75	1.15	1.69	2.57	4.33

	Correlation (r)	True Odds Ratio			
	TCM at A03 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
DCAN at site A03	0.80	1.16	1.75	2.74	4.78
DCAA at site A03	0.62	1.12	1.54	2.18	3.35
TCAA at site A03	0.69	1.13	1.62	2.38	3.84

	Correlation (r)	True Odds Ratio			
	TCM at A04 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
DCAN at site A04	0.81	1.16	1.75	2.76	4.84
DCAA at site A04	0.68	1.13	1.60	2.34	3.73
TCAA at site A04	0.80	1.16	1.74	2.73	4.75

Table 6-2. Summary of high and low r-values for investigating measured causal agent

Location and "measured" compound	**High r**			**Low r**		
	r 95% CI	**Sample size (n)**	**"True" Compound**	**r 95% CI**	**Sample size (n)**	**"True" Compound**
Daily						
City A TCM WTP1	0.92 0.83,0.96	27	DCAN	0.79 0.59,0.90	27	DCAA
City A TCM WTP2	0.86 0.71,0.93	27	DCAN	0.55 0.20,0.78	25	DCAA
City A TCM DS	0.84 0.68,0.92	29	DCAN	0.62 0.33,0.80	29	DCAA
City B TCM WTP	0.74 0.51,0.87	29	BDCM	0.22 -0.16,0.54	29	TCAA
City B TCM DS	0.91 0.82,0.96	29	BDCM	0.01 -0.36,0.38	28	DCAA
Monthly*						
City A TCM WTP1	0.86 0.76,0.92	48	BDCM	0.82 0.68,0.90	70	DCAA
City A TCM WTP2	0.90 0.83,0.94	48	BDCM	0.73 0.54,0.85	39	TCAA

WTP = water treatment plant; DS = distribution system

*In the monthly data, TCM and TTHM provided the highest correlations with each other: r=1.00 for all TCM/TTHM correlations. This is to be expected since TCM provides the major contribution to TTHM in all three systems. It was deemed more useful, therefore, to look at the second highest r-value when investigating these monthly values.

†City C does not have a water treatment plant, therefore all sampling points are in the distribution system

Table 6-2. continued, Summary of high and low r-values for investigating measured causal agent

Location and "measured" compound	High r			Low r		
	r 95% CI	Sample size (n)	"True" Compound	r 95% CI	Sample size (n)	"True" Compound
City A TTHM WTP1	0.83 0.71,0.90	48	BDCM	0.77 0.61,0.87	43	TCAA
City A TTHM WTP2	0.83 0.71,0.90	48	BDCM	0.59 0.35,0.76	42	TCAA
City B TCM WTP	0.96 0.92,0.98	36	DCAA	0.86 0.68,0.94	21	BDCM
City B TCM DS	0.96 0.88,0.99	15	THAA	0.58 0.14,0.83	17	BDCM
City B TTHM WTP	0.96 0.90,0.98	21	DCAA	0.85 0.69,0.93	27	BDCM
City B TTHM DS	0.94 0.87,0.97	25	THAA	0.49 0.11,0.75	24	BDCM
City C TCM †	0.99 0.97,1.00	17	BDCM	-0.01 -0.47,0.46	18	TCAA
City C TTHM †	0.99 0.91,1.00	6	BDCM	-0.01 -0.49,0.47	17	TCAA

WTP = water treatment plant; DS = distribution system

*In the monthly data, TCM and TTHM provided the highest correlations with each other: r=1.00 for all TCM/TTHM correlations. This is to be expected since TCM provides the major contribution to TTHM in all three systems. It was deemed more useful, therefore, to look at the second highest r-value when investigating these monthly values.

†City C does not have a water treatment plant, therefore all sampling points are in the distribution system

For daily sampling, the non-HAA compounds (DCAN and BDCM) correlate better with TCM than do the HAAs, as expected. Correlations are generally higher between TCM and other DBP species in the City A daily data than the City B daily data. This confirms the expected behaviour in a chloraminated system (City A) where the DBP concentrations are likely to vary less with time in the distribution system than in a chlorinated system (City B). As expected, BDCM, a trihalomethane, has the highest correlation (aside from TCM and TTHM correlated with each other) when correlated with both TCM and TTHM for both the City A monthly data and the City C #2 and #3 monthly data. TCAA and DCAA have the lowest correlations when correlated with TCM and TTHM in the City A and City C #2 and #3 monthly data. The City B monthly data exhibit the opposite behaviour. With the exception of the almost-perfect correlations between TCM and TTHM, the high r-values are exhibited by the HAAs, while the low r-values occur with BDCM, which is contrary to expectations. City C monthly data exhibit some negative r-values, the implications of which were discussed in Chapter 5.

From the correlations between TCM/TTHM and other DBPs, we can infer the varying degrees of exposure misclassification if TCM or TTHM data are used as surrogate measures for other compounds that might be causal agents. What we ultimately need to know is the effect this misclassification will have on the OR of a study if the TCM/TTHM concentrations are used as the measure of exposure when in fact another DBP compound is the true causal agent. For the purposes of the analysis method, we will assume that the other DBP compounds represent the "true" data sets and the TCM/TTHM data represent the "measured" data sets. Measured ORs have been calculated for the TCM/TTHM data for four "true" ORs: 1.20, 2.00, 3.50, and 7.00. For the purposes of this exercise, the "true" ORs are considered to be the result if the "true" other DBP data were used as the exposure measures. The measured ORs were calculated using the method outlined in Chapter 3 and briefly reviewed in Chapter 5. Tables with the calculated ORs are presented in the Appendix for Chapter 6. The negative correlations in the City C data result in measured ORs below one, as would be expected from Figures 5-1 and 5-2. All r-values less than one ($r<1.00$) result in varying degrees of OR attenuation, following Figures 5-1 and 5-2 and as seen in the tables in the appendix.

Having seen the potential for exposure misclassification as a result of measuring alternate DBPs and the subsequent effect on OR attenuation, we can turn our attention to the main question we are trying to answer. The question "How good does exposure assessment have to be?" can be re-worded to look at the effect of measuring alternate potential causal agents. The new question is: "What effect on the OR will measuring alternate potential causal agents have?". To answer this question, we look at how many data set pairs fall within 10% of the "true" OR. Given this parameter, we can calculate the level of correlation required between the measured data and the "true" data. This correlation value depends on the "true" OR. For OR_T of 7.00, r-values of 0.95 or greater are required for the measured OR to be within 10% of the OR_T. For OR_T of 3.50, r-values of 0.92 or greater are required. For OR_T of 2.00, r-values of 0.85 or greater are required. For $OR_T = 1.2$, r-values of only 0.42 or greater are required. As the OR_T increases, the measured data must be correlated more closely to the "true" data to produce a measured OR within 10% of the "true" OR. A summary of the number of measured data sets that result in the required r-values for each city and sampling frequency is presented in Table 6-3 and Figures 6-1 to 6-3.

Table 6-3. Number of r-values within the required range

	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00	Total # of data sets
	# data sets with r≥0.42	# data sets with r≥0.85	# data sets with r≥0.92	# data sets with r≥0.95	
Daily					
City A TCM WTP	6	3	1	0	6
City A TCM DS	12	0	0	0	12
City B TCM WTP	1	0	0	0	3
City B TCM DS	8	1	0	0	12
Monthly					
City A TCM WTP	8	7	0	0	8
City A TTHM WTP	8	0	0	0	8
City B TCM WTP	4	4	3	2	4
City B TCM DS	16	11	9	3	16
City B TTHM WTP	4	4	3	3	4
City B TTHM DS	15	3	2	0	16
City C TCM sampling location #2*	4	1	0	0	4
City C TCM DS**	16	10	4	4	52
City C TTHM sampling location #2*	4	2	1	1	4
City C TTHM DS**	15	12	2	1	52

* before the open reservoir and additional chlorine

** after the open reservoir and additional chlorine

Fig 6-1. Percent of data set pairs with measured ORs within 10% of the true OR (OR_T) for daily sampling, continuous treatment

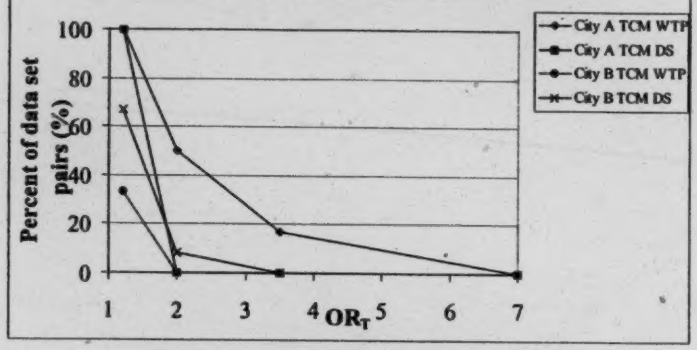


Fig 6-2. Percent of data set pairs with measured ORs within 10% of the true OR (OR_T) for monthly sampling, continuous treatment

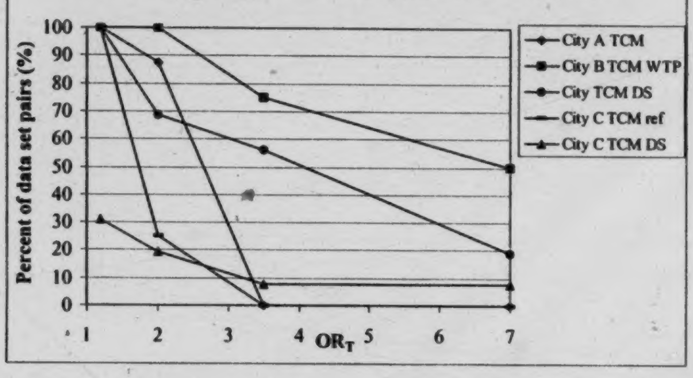
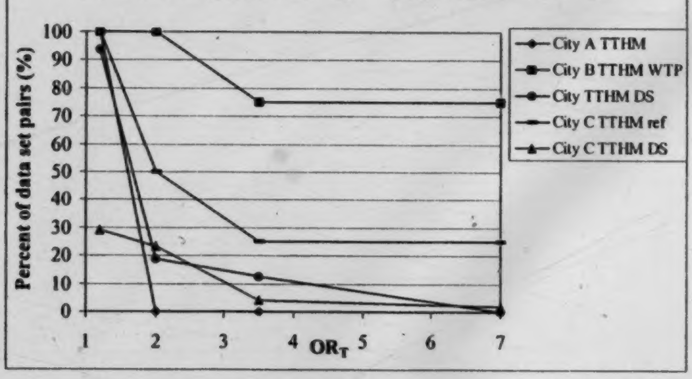


Fig 6-3. Percent of data set pairs with measured ORs within 10% of the true OR (OR_T) for monthly sampling, continuous treatment



There is evidence of differences between the chloraminated system and the chlorinated system. At the lowest true OR (OR_T=1.20) in the daily data for the chloraminated system (City A), TCM is a better surrogate measure of the three other DBP species than daily TCM data in the chlorinated system (City B). All (100%) of the daily data sets in the chloraminated system fall within the critical range at OR_T=1.20. Only one third (33%) of the daily water treatment plant data sets and two thirds (67%) of the daily distribution system data sets in the chlorinated system fall within the critical range at this OR_T. In the chlorinated system, daily TCM distribution system data can only be used consistently to predict BDCM, and only for the lowest OR_Ts. At higher OR_Ts, daily TCM data in the chlorinated system are not good surrogate measures for other DBPs. Two thirds (67%) of the daily water treatment plant data sets for the chloraminated system fall within the critical range at OR_T=2.00. However, the distribution system data sets do not follow suit at this OR_T, and no data set pairs fall into the critical range at higher OR_Ts.

Monthly TCM and TTHM data in the chloraminated system can be used consistently as surrogate measures for all other DBP compounds measured at OR_T=1.20, but less so at higher OR_Ts. Only water treatment plant data were used because there were no monthly distribution system data available for DBPs besides TTHM.

At OR_T=1.20, monthly TCM and TTHM data in the chlorinated system in both the water treatment plant and the distribution system provide a good surrogate measure for all other measured DBPs. At higher OR_Ts, TCM is best used as a surrogate measure for some HAAs at some locations, but not for BDCM.

As expected, these data confirm that chloraminated systems are more forgiving than chlorinated systems when using TCM or TTHM as surrogate measure for other DBPs.

City C is a unique case of a chlorinated system for several reasons, one being the inclusion of an open reservoir in the distribution system between the first and second chlorination points. At OR_T=1.20, the City C monthly TCM and TTHM data prove to be good surrogate measures for all other measured compounds before the open reservoir. After the open reservoir and at higher OR_Ts before the open reservoir, TCM and TTHM are best used as surrogates for BDCM.

The data in this chapter confirm that if the OR_T is expected to be high measuring TCM or TTHM as surrogates of other DBPs will likely result in substantial exposure misclassification. This observation is more critical to chlorinated systems than chloraminated systems, particularly for daily sampling. The exception to this observation is the City B monthly water treatment plant data, where both TCM and TTHM provide an adequate surrogate measure for some HAAs, even at the higher OR_Ts.

Application of method to literature values

Other researchers have investigated the relationship between DBP species. In one study (Keegan et al., 2001), routine monitoring data (mostly monthly samples at random sites in the distribution system) over a 5-year period were used to determine a relationship between THM species. The four individual THMs were measured (TCM, BDCM, DBCM, TBM) with the sum of the four equaling the total THM (TTHM) measurement. The data were collected from 288 water supply zones in the north west of England. The majority of the raw water was from surface sources. The type of disinfection (chlorine, chloramine) was not mentioned. TTHM concentrations were correlated with TCM, BDCM, and DBCM concentrations. TBM was not included in these analyses because a large number of the TBM measurements were below the detection limit. The correlations between TTHM and BDCM, and TTHM and TCM were $r=0.62$ and $r=0.98$, respectively. These r-values were significant at $p<0.01$. The correlation between TTHM and DBCM was $r=-0.09$ and was not significant at $p<0.01$.

In another study (Singer et al., 1995), the association between THMs and HAAs was investigated. Samples from six North Carolina utilities were collected at the water treatment plants and two or three distribution system locations. Sampling was conducted three times over a year-long period. The samples were analyzed for THMs and HAAs, among other parameters. Only chlorine was used as a disinfectant by the utilities in this study and the raw water was from surface sources. The correlation between TTHM and THAA concentrations was found to be $r=0.897$ (95%CI 0.849,0.930).

The method introduced in Chapter 3 and employed in Chapters 5 and 6 can also be applied to these data. Table 6-4 shows the r-values and resulting attenuated ORs for each of the four "true" ORs. In employing the method of calculating OR attenuation, TTHM is considered the "measured" value and BDCM, DBCM, TCM and THAA are each considered the "true" values.

Table 6-4. OR attenuation method applied to correlations between DBP species from the literature

Singer et al., 1995					
	Correlation (r)	True Odds Ratio			
	TTHM (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
THAA	0.90	1.10	1.86	3.07	5.72
Keegan et al., 2001					
	Correlation (r)	True Odds Ratio			
	TTHM	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
TCM	0.98	1.20	1.97	3.41	6.73
BDCM	0.62	1.12	1.54	2.17	3.34
DBCM	-0.09	0.98	0.89	0.80	0.71

OR_T = the assumed OR for the true data
OR_X = the calculated OR for the measured data = OR_T

From the Keegan et al. (2001) study, measuring TTHM in lieu of DBCM would result in extreme exposure misclassification and highly attenuated ORs. All the other r-values result in measured ORs within 10% of the "true" OR at OR_T=1.20. In addition, THAA in the Singer study resulted in measured ORs within 10% of the "true" OR at OR_T up to 2.00. This supports the observation in this chapter that at longer sampling frequencies (monthly or longer), TTHMs can provide adequate surrogate measures of THAAs, up to certain OR_Ts in chlorinated systems. The TCM correlation in the Keegan study resulted in measured ORs within 10% of the "true" OR at all OR_Ts. This supports the finding in this chapter that TTHM and TCM are highly correlated and can be substituted as surrogate measures of each other without much exposure misclassification. The BDCM correlations and subsequent attenuated ORs were within the range of r-values found for correlations between BDCM and TTHM in this study for monthly sampling in City B.

Categorical analysis

Many epidemiology studies group exposure data into categories rather than using continuous data. This is often done to account for imprecisions in the continuous data from sampling and analysis. However, the differences in exposure misclassification between categorical and continuous treatment of exposure data may result in different levels of OR attenuation.

All the exposure data in these analyses were divided into quartiles. The top quartile was designated as "exposed" while the lower three quartiles were designated as "not exposed". As with the continuous treatment, "true" and "measured" data were required for the categorical treatment. The TCM and

TTHM data sets were designated as the "measured" data, and the other DBP data sets were designated as the "true" data, as with the continuous treatment. Sensitivity and specificity were calculated for each pair of measured/true data sets. Examples of the calculation of sensitivity and specificity are provided in Chapter 1. In Chapter 3 we introduced equations that link exposure misclassification, as measured by sensitivity and specificity, to OR attenuation. The equations were reviewed in Chapter 5. Table 6-5 is an example of the calculated sensitivities, specificities, and attenuated ORs from City A daily data. The rest of the data are presented in the Appendix.

Table 6-5. Sensitivity and specificity between TCM and other DBP concentrations at water treatment plants 1 and 2 and distribution system locations in City A, and their effects on the odds ratio

City A Daily						
			True Odds Ratio			
	TCM at WTP1 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
DCAN at WTP1	1.00	0.90	1.14	1.72	2.81	5.34
DCAA at WTP1	0.71	0.85	1.09	1.43	1.98	2.91
TCAA at WTP1	0.67	1.00	1.18	1.86	2.94	4.79
			True Odds Ratio			
	TCM at WTP2 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
DCAN at WTP2	0.39	0.89	1.06	1.28	1.59	2.02
DCAA at WTP2	0.67	0.74	1.06	1.26	1.57	2.07
TCAA at WTP2	0.75	0.71	1.06	1.29	1.66	2.28
			True Odds Ratio			
	TCM at A01 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
DCAN at site A01	0.80	0.96	1.15	1.74	2.73	4.60
DCAA at site A01	0.43	0.91	1.08	1.36	1.76	2.34
TCAA at site A01	0.42	1.00	1.17	1.77	2.66	3.97
			True Odds Ratio			
	TCM at A02 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
DCAN at site A02	0.50	0.96	1.12	1.56	2.21	3.22
DCAA at site A02	0.40	0.92	1.08	1.35	1.75	2.32
TCAA at site A02	0.50	1.00	1.17	1.80	2.74	4.20
			True Odds Ratio			
	TCM at A03 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
DCAN at site A03	1.00	0.68	1.09	1.44	2.10	3.64
DCAA at site A03	0.86	0.64	1.06	1.30	1.70	2.45
TCAA at site A03	0.67	0.65	1.04	1.19	1.40	1.74
			True Odds Ratio			
	TCM at A04 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
DCAN at site A04	0.67	0.95	1.14	1.64	2.43	3.78
DCAA at site A04	0.75	0.84	1.09	1.44	2.00	2.99
TCAA at site A04	0.64	1.00	1.18	1.85	2.90	4.67

The categorical treatment shows similar trends to the continuous treatment. In the daily sampling, the non-HAA compounds generally correlated better with TCM than did the HAAs. In addition, less OR attenuation was seen between TCM and other DBPs in the chloraminated system (City A) than in the chlorinated system (City B). In the monthly sampling, BDCM showed similar trends when correlated with either TCM or TTHM in City A, while City B showed no consistent pattern in OR attenuation between BDCM and HAAs when correlated with either TCM or TTHM. As in Chapter 5, the categorical treatment resulted in lower measured ORs than the continuous treatment for corresponding data set pairs.

The continuous treatment of the City C monthly data suggested that BDCM compared with either TTHM or TCM would result in the highest measured ORs. This is also observed in the categorical treatment.

The purpose of this chapter has been to answer another facet of the question: "How good does exposure assessment have to be?". We have seen in the data the extent of exposure misclassification and the subsequent effect of that misclassification on the OR of measuring TCM or TTHM rather than a more toxicologically plausible causal agent. In the continuous data treatment, we determined the correlation required between the "true" and "measured" data sets to result in an OR_X within 10% of the OR_T. However, with the categorical analysis, we have two independent variables affecting the OR_X. As a result, it is difficult to produce a simple table to describe the sensitivity and specificity requirements to produce an OR_X within 10% of the OR_T. Therefore, Table 6-6 and Figures 6-4 to 6-6 show the number of data sets within each city and sampling frequency that result in the OR_X within the required range, based only on the final OR_X and not on the sensitivity and specificity.

Table 6-6. Number of measured ORs within 10% of the true OR

Location	**OR_T=1.20**	**OR_T=2.00**	**OR_T=3.50**	**OR_T=7.00**	
	# of data sets with $OR_X \geq 1.08$	**# of data sets with $OR_X \geq 1.80$**	**# of data sets with $OR_X \geq 3.15$**	**# of data sets with $OR_X \geq 6.30$**	**Total # data sets**
Daily					
City A TCM WTP	5	0	0	0	6
City A TCM DS	11	0	0	0	12
City B TCM WTP	1	0	0	0	3
City B TCM DS	4	1	1	1	12
Monthly					
City A TCM WTP	7	0	0	0	8
City A TTHM WTP	8	1	0	0	8
City B TCM WTP	1	0	0	0	4
City B TCM DS	10	0	0	0	6
City B TTHM WTP	4	0	0	0	4
City B TTHM DS	15	0	0	0	6
City C TCM reference location #2	4	2	0	0	4
City C TCM DS	15	2	2	2	52
City C TTHM reference location #2	4	1	1	1	4
City C TTHM DS	14	1	1	1	52

[1] City C data were analyzed using both sampling points #2 and #3 as reference points for comparison

WTP2=Water treatment plant #1; WTP1 = Water treatment plant #2

Fig 6-4. Percent of data set pairs with measured ORs within 10% of the true OR (OR_T) for daily sampling, categorical treatment

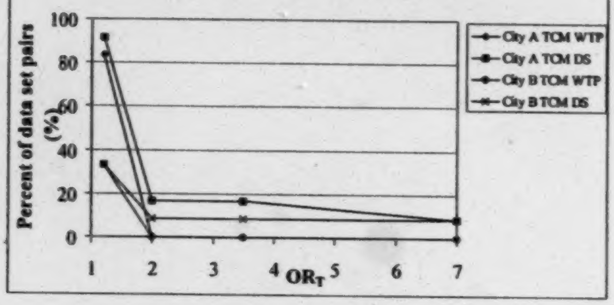


Fig 6-5. Percent of data set pairs with measured ORs within 10% of the true OR (OR_T) for monthly sampling, categorical treatment

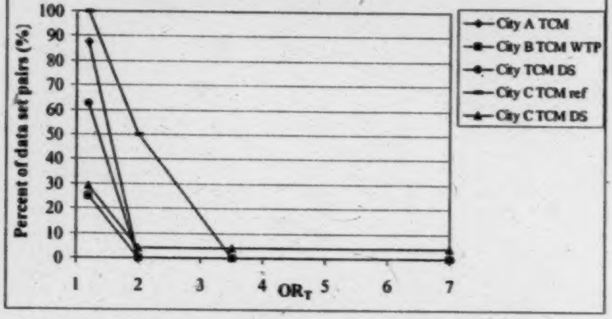
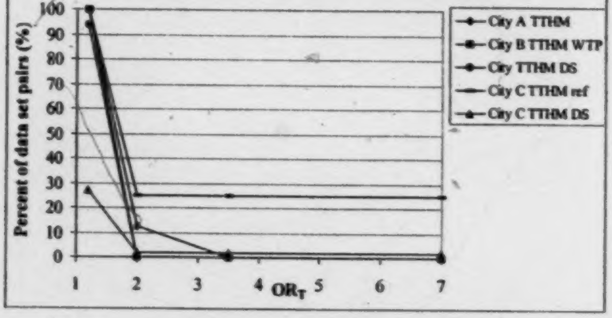


Fig 6-6. Percent of data set pairs with measured ORs within 10% of the true OR (OR_T) for monthly sampling, categorical treatment



We observe similar trends in terms of numbers of data sets producing measured ORs within 10% of the OR_T as in the continuous table. However, fewer overall data sets in the categorical treatment are within the 10% range. There are noticeably fewer categorical data sets in the 10% range at OR_Ts higher than 1.20 than in the continuous treatment. This could be a result of the overall lower measured ORs in the categorical treatment compare to the continuous treatment. This phenomenon will be explained further in Chapter 8. However, the categorical data generally support the trends seen in the continuous data.

Summary and conclusions

In this chapter we have investigated the potential for exposure misclassification when TCM or TTHM are measured as the causal agents for exposure rather than other, toxicologically more plausible, causal agents. The expected result that TCM and TTHM are good surrogate measures of THM and DCAN was confirmed by most of the daily and monthly data. A notable exception to this observation occurred in the City B monthly data, where TCM and TTHM were better surrogate measures of the HAAs than of BDCM for both the categorical and continuous treatment.

At the lowest OR_T, (OR_T=1.20) for both the categorical and continuous treatment, TCM and TTHM seem to provide adequate surrogate measures of the other DBP species. In addition, this trend was stronger in the chloraminated system than in the chlorinated system. As the OR_T increases, however, the utility of TCM and TTHM as surrogate measure for the other DBP species decreases. Therefore, at low OR_Ts these results suggest that measuring TCM or TTHM as a surrogate for other more toxicologically plausible causal agents will not result in substantial exposure misclassification and subsequent attenuated ORs, particularly in chloraminated systems over chlorinated systems. However, as the OR_T increases, the chance of misclassification also increases and the need to measure the actual causal agent rather than a surrogate becomes clear. This observation is more critical to chlorinated systems than chloraminated systems, particularly for daily sampling.

References

Keegan, T., H. Whitaker, M.J. Nieuwenhuijsen, M.B. Toledano, P. Elliott, J. Fawell, M. Wilkinson, N. Best. 2001. Use of Routinely Collected Data on Trihalomethane in Drinking Water for Epidemiological Purposes. *Occupational and Environmental Medicine* **58**: 447-452.

Singer, P.C., A. Obolensky, A. Greiner. 1995. DBPs in Chlorinated North Carolina Drinking Waters. *Journal of the American Water Works Association* **87**(10): 83-92.

Chapter 7

Potential for Exposure Misclassification and Subsequent OR Attenuation as a Result of Monthly Convenience Sampling

Introduction

In Chapter 5 we discussed the requirements for a comprehensive version of exposure assessment in the absence of biomarkers of exposure for epidemiology studies investigating a possible causal association between exposure to chlorinated disinfection by-products in drinking water and adverse reproductive and developmental effects. One of the criteria for this comprehensive exposure assessment was continuous monitoring of disinfection by-product (DBP) concentrations at the points of exposure. In Chapter 5 we discussed the potential for exposure misclassification resulting from sampling at a distance from the points of exposure. In Chapter 6 we discussed the potential for exposure misclassification resulting from measuring TCM and TTHM rather than other more toxicologically feasible causal agents. In this chapter we will investigate the potential for exposure misclassification as a result of convenience sampling on different days of the month or using summary variables such as the mean, maximum, or minimum monthly concentrations.

As discussed in Chapter 2, many of the epidemiology studies to date that have used DBP concentrations as a measure of exposure to DBPs have relied on monthly monitoring samples at water treatment plants or in the distribution system. While continuous monitoring of DBP concentrations at all points of exposure is likely out of the realm of practicality, daily samples at the water treatment plants are certainly possible. Daily water treatment plant samples have the inherent problem of concentration variations as a result of distance from the points of exposure. However, they provide a means of determining the effect of using convenience sampling or summary variables as measures of exposure. As seen in Chapter 5, the extent of misclassification caused by sampling away from the points of exposure depends on many factors. In some cases, water treatment plant data are a valid approximation of distribution system samples for the purposes of exposure assessment in these epidemiology studies in that they are not predicted to result in substantial exposure misclassification. In addition, daily sampling at the point of exposure would present a substantial burden for the study subjects. In this chapter, daily TTHM and TCM monitoring data from the two water treatment plants in City A will be used. Daily monitoring data were not available for the City A distribution system, or from other cities and for other DBP species.

The ultimate question being asked, then, is: "Should daily sampling replace monthly sampling at water treatment plants and in the distribution system to account for DBP concentration variations as they relate to exposure to DBPs in drinking water?" Monthly sampling at most water utilities tends to be convenience sampling in the sense that a convenient day of the month is chosen as the sampling day, regardless of the maximum, minimum or mean DBP concentrations during the month. Accordingly, certain sub-questions must be asked. Can monthly convenience sampling account for DBP concentration variations during the month? Is exposure misclassification likely to occur using convenience sampling rather than monthly mean, maximum, or minimum DBP concentrations? What effect is such misclassification likely to have on the odds ratio (OR) of epidemiology studies? If daily sampling at the water treatment plant is implemented, should all the data be used in epidemiologic exposure assessment, or will mean monthly values calculated from the daily data suffice? Again to what extent will exposure misclassification and OR attenuation occur? In this chapter, we will attempt to answer these questions.

Daily TTHM and TCM data were available from the two water treatment plants in City A for the years 1997-2000. For each month, the maximum, minimum, and mean values were calculated. In addition, four convenience sampling days were chosen. The four convenience sampling days were chosen to be: 1) the second Tuesday of every month, 2) the 18th of every month, 3) the first Wednesday of every month, and 4) the 4th of every month. These days were chosen to provide a representation of convenience monitoring that could occur at water utilities. Each of the four convenience sampling days generated a data set that could be compared to the mean, minimum, and maximum values. All data were treated both continuously and categorically. The extent of misclassification was measured by the correlation coefficient between a "true" and "measured" data set for the continuous treatment, and by sensitivity and specificity between the "true" and "measured" data sets for the categorical treatment. These measures were then utilized in the method discussed in Chapter 3 and reviewed in Chapter 5 to quantify the attenuation of the OR as a result of using the "measured" data to calculate the OR rather than the "true" data. In order to compare two data sets categorically or continuously using the methods discussed in Chapter 3, we must term one data set "measured" and other "true". In comparing convenience sampling with maximum, minimum, and mean values, we designate the convenience data as "measured" and the other data sets "true". In comparing the mean DBP concentrations with the maximum, minimum, and random values, we designate the mean data as "measured" and the other data sets "true". The words "true" and "measured" are used purely to keep the semantics associated with these statistical methods consistent with Chapters 5 and 6. None of the data sets can be reliably termed "true" in the sense of being a better measure of individual exposure to DBPs for subjects in an epidemiology study. The correlation coefficients, sensitivity, specificity, and resulting attenuated ORs are presented in the Appendix.

Analysis of continuous exposure data

First we examine whether monthly convenience sampling can account for variations in DBP concentrations. Since the causal agents and harmful levels of exposure to the causal agents have not been identified, we are particularly interested in the ability of convenience sampling to account for the maximum monthly concentrations. Convenience sampling TCM and TTHM data were compared to maximum, minimum, and mean TCM and TTHM concentrations at each water treatment plant and the correlation coefficients were calculated. Table 7-1 shows an example of the correlations and attenuated ORs. The rest of the data are in the Appendix.

Table 7-1. Water treatment plant #1 TCM data in City A; correlations between the "measured" convenience sampling data and the "true" data with the resulting OR attenuation due to exposure misclassification

Water treatment plant #1 TCM data

"True" data	Correlation (r)	True Odds Ratio			
	Convenience sampling #1 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_{Xa}	OR_{Xa}	OR_{Xa}	OR_{Xa}
Maximum TCM	0.89	1.18	1.86	3.06	5.69
Minimum TCM	0.89	1.18	1.85	3.05	5.66
Mean TCM	0.93	1.19	1.91	3.23	6.17

"True" data	Correlation (r)	True Odds Ratio			
	Convenience sampling #2 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_{Xa}	OR_{Xa}	OR_{Xa}	OR_{Xa}
Maximum TCM	0.90	1.18	1.86	3.08	5.74
Minimum TCM	0.92	1.18	1.89	3.17	6.01
Mean TCM	0.95	1.19	1.93	3.27	6.30

"True" data	Correlation (r)	True Odds Ratio			
	Convenience sampling #3 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_{Xa}	OR_{Xa}	OR_{Xa}	OR_{Xa}
Maximum TCM	0.87	1.17	1.83	2.97	5.43
Minimum TCM	0.82	1.16	1.77	2.80	4.95
Mean TCM	0.87	1.17	1.83	2.98	5.46

"True" data	Correlation (r)	True Odds Ratio			
	Convenience sampling #4 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_{Xa}	OR_{Xa}	OR_{Xa}	OR_{Xa}
Maximum TCM	0.88	1.17	1.84	3.00	5.51
Minimum TCM	0.84	1.16	1.78	2.85	5.08
Mean TCM	0.88	1.17	1.84	3.00	5.52

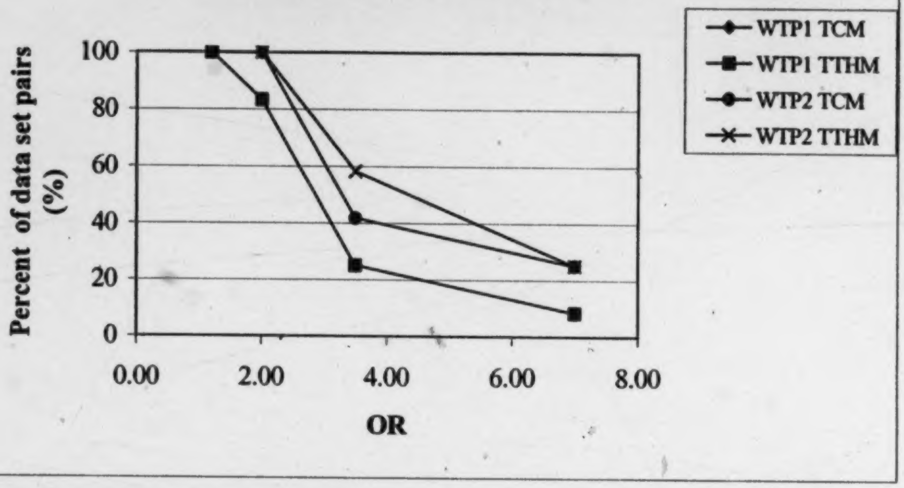
The correlations between the convenience sampling data sets and the mean, maximum, and minimum data are all high. Any differences between the water treatment plants are well within the variation expected from the slightly different water treatment methods and differences in source water intake locations between the two plants. In addition, the results of the TCM and TTHM data are very similar, which is to be expected because TCM makes up the majority of TTHM. Additionally, the results when compared between convenience data sets are remarkably similar, suggesting that the day of convenience sampling has little effect on the correlations with maximum, minimum, and mean values. As a general trend, the convenience sampling data sets produce the highest correlations when compared with the mean data. The lowest correlations are almost evenly split between the maximum and minimum data.

In order to determine the utility of convenience sampling for exposure assessment, we need to know the effect of any misclassification on the OR of an epidemiology study. For this reason we are interested in the number of data sets that result in a "measured" OR that is within 10% of the "true" OR. Table 7-2 and Figure 7-1 show the number of data set pairs that fulfill these conditions.

Table 7-2. Number of data set pairs with correlation coefficients within the required range for the OR_X to be within 10% of the OR_T

	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00	Total # of data set pairs
	# data sets with r≥0.42	# data sets with r≥0.85	# data sets with r≥0.92	# data sets with r≥0.95	
Water treatment plant #1 TCM	12	10	3	1	12
Water treatment plant #1 TTHM	12	10	3	1	12
WTP2 TCM	12	12	5	3	12
WTP2 TTHM	12	12	7	3	12

Fig 7-1. Percent of data set pairs with measured ORs within 10% of the true OR (OR_T), continuous treatment



The results in Table 7-2 and Figure 7-1 show that very little OR attenuation occurs up to OR_T=2.00. This suggests that at lower OR_Ts, convenience sampling can provide a reasonable estimation of maximum, minimum, and mean monthly values at the water treatment plants.

In addition to monthly convenience sampling, we are interested in the correlations between mean monthly concentrations and maximum or minimum monthly values. If sampling at the water treatment plant were done daily and the data used for epidemiologic exposure assessment, mean concentrations may be preferred for exposure assessment over maximum concentrations because mean concentrations may be more representative of DBP concentrations over the entire month, while maximum concentrations may be unusual occurrences. Median concentrations were originally included in this analysis; however, the mean and median data sets had correlations of 0.99 or 1.00, resulting in little or no OR attenuation. In addition, the correlations between the median and the maximum and minimum data sets were the same as the correlations between the mean and the maximum and minimum data sets. Therefore, the results for the median data set are not included here. The correlations and the resulting ORs are presented in the Appendix. The correlations with the convenience sampling data sets are included in the tables for comparison.

The maximum and minimum data are highly correlated with the mean data. All but one of the correlations for the maximum and minimum data sets fall within the range of OR attenuation less than 10% of the OR_T for all the OR_Ts examined. There is one exception: at OR_T=7.00, the minimum TCM values at water treatment plant #2 do not fall within 10% of the "true" OR. These results suggest that very little exposure misclassification and OR attenuation would occur if mean

DBP concentrations were used in place of maximum or minimum values. There are some caveats associated with these results. These statistical methods account for OR attenuation caused by relative variations in data sets, not the absolute differences in DBP concentrations between data sets. It would be difficult to predict the effect on the OR in a real study because the causal agents and critical exposure concentrations for those causal agents are not known. However, these results look promising for the use of monthly monitoring data from chloraminated systems in epidemiology studies.

Analysis of categorical exposure data

The data were divided into categories as described in Chapter 5. Table 7-3 shows an example of the calculated sensitivities, specificities, and attenuated ORs. The rest of the data are presented in the Appendix for Chapter 7.

Table 7-3. Categorical treatment of City A data comparing the "measured" convenience sampling data with the "true" values and the resulting effect on the OR

Water treatment plant #1 TCM						
			True Odds Ratio			
	Convenience #1 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True" data	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
Maximum TCM	0.50	1.00	1.17	1.80	2.74	4.20
Minimum TCM	0.29	1.00	1.16	1.74	2.54	3.67
Mean TCM	0.33	1.00	1.17	1.75	2.58	3.77
	Convenience #2 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True" data	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
Maximum TCM	0.43	0.98	1.13	1.61	2.30	3.33
Minimum TCM	0.43	0.98	1.13	1.61	2.31	3.35
Mean TCM	0.50	0.98	1.14	1.65	2.41	3.59
	Convenience #3 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True" data	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
Maximum TCM	0.43	1.00	1.17	1.78	2.67	4.00
Minimum TCM	0.29	0.98	1.11	1.51	2.06	2.84
Mean TCM	0.33	0.98	1.12	1.55	2.16	3.02
	Convenience #4 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True" data	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
Maximum TCM	0.43	1.00	1.17	1.78	2.67	4.00
Minimum TCM	0.29	0.98	1.11	1.51	2.06	2.84
Mean TCM	0.43	0.98	1.13	1.61	2.31	3.35

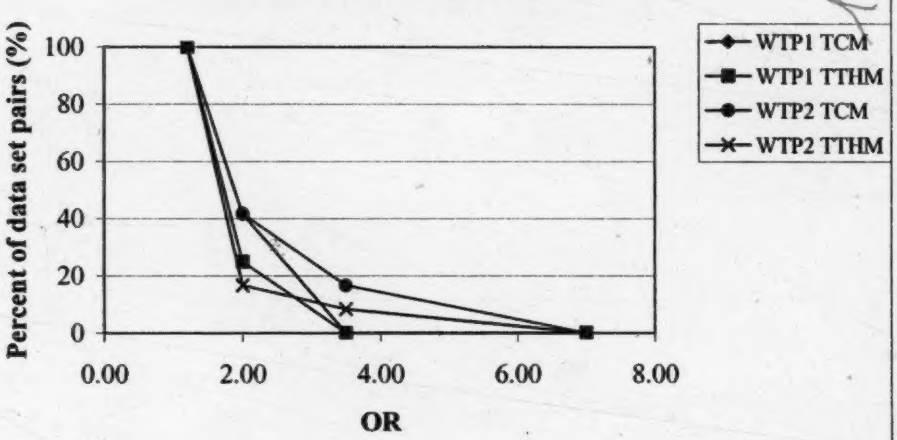
The categorical data show the same trends as the continuous data. However, the measured ORs in the categorical treatment are generally lower than those in the continuous treatment. The results between the convenience sampling data sets and between the two water treatment plants are very similar. As expected, the TCM and TTHM results are similar.

We are ultimately interested in the effect of the exposure misclassification on the OR. To this end, we can quantify the number of data set pairs that result in a "measured" OR attenuated within 10% of the "true" OR. Table 7-4 and Figure 7-2 show this summary.

Table 7-4. Number of data set pairs with measured ORs within 10% of the true OR

Location	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00	
	# of data sets with $OR_X \geq 1.08$	# of data sets with $OR_X \geq 1.80$	# of data sets with $OR_X \geq 3.15$	# of data sets with $OR_X \geq 6.30$	Total # data sets
WTP1 TCM	12	5	0	0	12
WTP2 TCM	12	3	0	0	12
WTP1TTHM	12	5	2	0	12
WTP2TTHM	12	2	2	0	12

Fig 7-2. Percent of data set pairs with measured ORs within 10% of the true OR (OR_T), categorical treatment



Again, the categorical data show similar trends to the continuous data. The main difference is that there are fewer data set pairs within 10% of the OR_T at higher "true" ORs in the categorical treatment than in the continuous treatment. This is a similar finding to Chapters 5 and 6. The choice of treating the data as continuous or categorical can have a substantial effect on the resulting apparent exposure misclassification and the resulting measured OR, particularly at higher OR_Ts. At the lowest OR_T, the difference is often negligible. This point will be discussed further in Chapter 8.

The second part of this discussion deals with the ability of monthly mean DBP concentrations to be used as measures of exposure as compared to monthly maximum and minimum concentrations. The results are presented in the Appendix.

The mean data compared with the maximum and minimum data show very similar OR attenuation for both TCM and TTHM. However, the measured categorical ORs are again lower than the corresponding measured ORs in the continuous treatment. While the mean DBP concentrations are an adequate surrogate of maximum and minimum DBP concentrations at $OR_T=1.20$ in the categorical treatment, they are not at higher OR_Ts. This is in direct contrast to the continuous results and show that the method of treating the data can affect the apparent exposure misclassification and the resulting measured OR quite substantially.

Conclusions

Several questions were addressed in this chapter. All the questions were aimed at determining whether monthly convenience sampling or monthly mean DBP concentrations adequately account for monthly variations in DBP concentrations and as a consequence provide an adequate measure of exposure for epidemiology studies without substantial exposure misclassification. The answers to these questions depend largely on whether the data were treated as continuous or categorical. The continuous treatment of both the convenience and mean data results in high, statistically significant correlations. In addition, almost all the data set pairs resulted in measured ORs within 10% of the OR_T at all OR_Ts for the mean values and at the lower OR_Ts for the convenience sampling. These results suggest that there is little exposure misclassification occurring between convenience sampling and the other measure at the lower OR_Ts and between mean DBP concentrations and maximum or minimum concentrations at all OR_Ts. Therefore the continuous results lead us to believe that convenience sampling at low OR_Ts and mean values at all OR_Ts can be used as measures of exposure with very little threat of exposure misclassification substantial enough to affect the measured OR. In addition, all convenience sampling data sets gave very similar results, suggesting that the day chosen for convenience sampling is irrelevant.

As in the continuous treatment, the categorical treatment showed similar results between water treatment plants, between TCM and TTHM, and between convenience sampling days. However, there is a major difference between the number of data set pairs resulting in measured ORs within 10% of the OR_T in the categorical treatment vs. the continuous treatment. Substantial numbers of data set pairs giving measured ORs within 10% of the OR_T were seen only in the lowest OR_T (OR_T=1.20) for both convenience and mean values. This observation suggests that in the categorical treatment, convenience sampling and mean values provide adequate surrogate measurements only at the very lowest OR_Ts. Therefore, the decision to use monthly convenience sampling or monthly mean values as the exposure measure will depend on the type of data treatment and the expected OR_T. The results of differences between continuous and categorical data treatment will be discussed in Chapter 8.

Chapter 8

Effects of categorical vs. continuous treatment of exposure data

In Chapters 5, 6, and 7, a trend was observed between the categorical and continuous treatments of the exposure data. In general, the categorical treatment followed the same trends as the continuous treatment. However, the categorical treatment resulted in measured ORs smaller than the measured ORs in the continuous treatment for corresponding data set pairs. This suggests that an epidemiology study using categorical data rather than continuous data would find lower associations between the exposure and the outcome. One reason for this discrepancy could lie in the cutpoints chosen for categorical treatment and their effect on sensitivity and specificity. To investigate this hypothesis, the City A daily data from Chapter 5 were re-analyzed categorically with cutpoints of tertiles, quartiles (original cutpoint), and sextiles. Halves were included for academic interest, to see the effect on sensitivity and specificity. Quartiles were used in the previous chapters because they are commonly used in epidemiology studies. The results are summarized in Tables 8-1 to 8-5, found at the end of this chapter.

Sensitivity and specificity are determined by the numbers in the cells of the 2x2 table. A refresher of sensitivity and specificity calculations may be useful at this point.

Table 8-6. Example of a 2x2 table to determine the sensitivity and specificity of a study

	True exposure		
Imperfect exposure	**Yes**	**No**	**Total**
Yes	100 (a)	5 (b)	105
No	10 (c)	85 (d)	95
Total	110	90	

Sensitivity is the proportion of those who are truly exposed who are measured as exposed:

$$\text{Sensitivity} = a / (a+c) = 100/110 = 0.91, \text{ or } 91\%$$

Specificity is the proportion of those who are truly not exposed who are measured as not exposed:

$$\text{Specificity} = d / (b+d) = 85/90 = 0.94, \text{ or } 94\%$$

In setting cutpoints, we are determining the number of subjects who are and who are not exposed, and are therefore influencing the numbers in the cells of the 2x2 table. In our analysis we set the highest category as the exposed category and all lower categories as not exposed. For example, in the quartiles, the top quartile was classified as "exposed" while the lower three quartiles were classified as "not exposed". In the sextiles, the top sextile was classified as "exposed" while the lower five sextiles were classified as "not exposed", and so on. As the number of categories increases, the number of subjects in each category decreases. For example, there will be fewer subjects in each sextile where the subjects are divided into six categories, than in each quartile where the subjects are divided into four categories. Therefore, as the number of categories increases (e.g. from tertiles to quartiles to sextiles), the number of exposed subjects will decrease and the number of not exposed subjects will increase in both the "true" and "measured" data sets. Remember that both the "true" and "measured" data sets were categorized the same way. In the halves cutpoint, half of our subjects were set as exposed and half as not exposed and we would expect to see very similar sensitivities and specificities for each data set pair. As expected for the halves cutpoint, sensitivity and specificity are very similar. The halves cutpoint will not be discussed further as they are unlikely to be used in "real life" epidemiology studies.

The main message from Tables 8-1 to 8-5 is that the different cutpoints can give very different results from each other and from the continuous analysis. One good example is DCAA A01. The continuous treatment suggests some misclassification and little OR attenuation, as does the tertiles cutpoint. The quartiles cutpoint, however, suggests perfect classification (no misclassification) and no OR attenuation. Alternately, the sextiles cutpoint shows a larger amount of misclassification than the continuous treatment or the tertiles cutpoint with greater resulting attenuation in the measured ORs. The effects are particularly noticeable at the larger "true" ORs.

Having noted the different results from various cutpoints and the continuous data, we can look for a trend. Perhaps sensitivity or specificity or OR attenuation show an increase or decrease as the number of categories increases? Unfortunately, there is no trend between cutpoints in the categorical treatment and r-values in the continuous treatment that is consistent for all the data. There are a few points where the r-value in the continuous treatment is low, and the sensitivity decreases and the specificity increases as the number of categories increases (e.g. TCMA03). However, this trend is not observed consistently in data with higher r-values in the continuous data. There is a general pattern that the categorical data result in lower ORs than the continuous data, although even this trend is not found in absolutely every case. The main observation from this analysis is that categorical treatment can result in widely varied results compared to continuous treatment. In addition, the cutpoint chosen for the categorical analysis has a large effect on the final ORs.

Another general trend observed in Chapters 5 through 7 in the quartiles categorical data is that the specificity is higher than the sensitivity. There are some exceptions to this observation, but the majority of the data support it. One possible explanation for this is a slight skewedness of the data distributions. The data are normally distributed for the most part, although a few data sets do exhibit a slight skewing. For the OR attenuation method to work, at least one data set of each pair must be normally distributed. To investigate the possible effect of skewedness on sensitivity and specificity calculations, the City A water treatment plant #1 daily TCM data were log transformed. The sensitivities and specificities as well as the attenuated ORs were re-calculated using these transformed data. The results show that there is little difference between the results of the log transformed data and the data in its original form. This suggests that the method is robust enough that skewedness to the degree found in these data do not affect the final results substantially. However, it does not explain the observation that most specificities in these analyses are greater than the corresponding sensitivities.

In a study for Health Canada (Reif et al., 2000), the authors observed that specificity has a greater effect on OR attenuation than sensitivity. Two examples of this were seen in these data as well. Table 8-7 shows these examples, as well as one example that does not support this observation. In these examples, either sensitivity or specificity is 1.00, so the effect of the other measure on OR attenuation can be clearly seen. In both these examples, the alternate (not 1.00) value of the sensitivity is the same as or lower than the alternate (not 1.00) value of the specificity. Therefore, if sensitivity and specificity have an equal effect on OR attenuation, we would expect to see lower measured ORs in the cases in this example where sensitivity is not equal to one than in the cases where specificity is not equal to one. However, the opposite is observed for the first two examples. The cases where specificity is "not 1.00" show greater OR attenuation than the cases where sensitivity is "not 1.00". This would seem to support the observation made by Reif et al. (2000) that specificity has a greater effect on OR attenuation than sensitivity.

Table 8-7. Three examples of the influence of specificity and sensitivity on OR attenuation

				OR_T= 1.20	**OR_T= 2.00**	**OR_T= 3.50**	**OR_T= 7.00**
	Sampling Location	**Sensitivity**	**Specificity**	**OR_X**	**OR_X**	**OR_X**	**OR_X**
Example #1							
City C#3	TCM5	1.00	0.75	1.10	1.50	2.25	4.00
City A monthly WTP2	TTHM4	0.75	1.00	1.19	1.89	3.05	5.17
Example #2							
City C#3	TTHM6	1.00	0.86	1.13	1.64	2.60	4.85
City A daily WTP2	DCAA A01	0.86	1.00	1.19	1.93	3.22	5.80
Example #3							
City C #2	TCAA8	1.00	0.93	1.16	1.78	2.95	5.69
City C #2	TCAA4	0.67	1.00	1.18	1.86	2.94	4.80

The first two examples support the observation of Reif et al. (2001). However, the third example contains an apparent anomaly. In this example, specificity has a greater effect than sensitivity up to $OR_T=2.00$. At $OR_T=3.50$ and higher, the opposite occurs. To find an explanation for this seeming anomaly, we started by plotting OR_X against OR_T for the anomalous sample and came up with Figure 8-1.

Figure 8-1 shows that the trend lines cross at a point just before $OR_T=3.50$. This suggests that specificity in this case has a greater effect on the "measured" ORs for "true" ORs up to 3.50. Above this "true" OR, the sensitivity has a greater effect than specificity. We plotted the other two examples (Figures 8-2 and 8-3) to see if this observation held true. Since the change-over point evidently comes at higher "true" ORs than we investigated in the main analysis, we included unrealistically high "true" ORs in the plot to see where the change-over point would come.

Figure 8-1. Anomalous example showing effect of sensitivity and specificity on OR_X

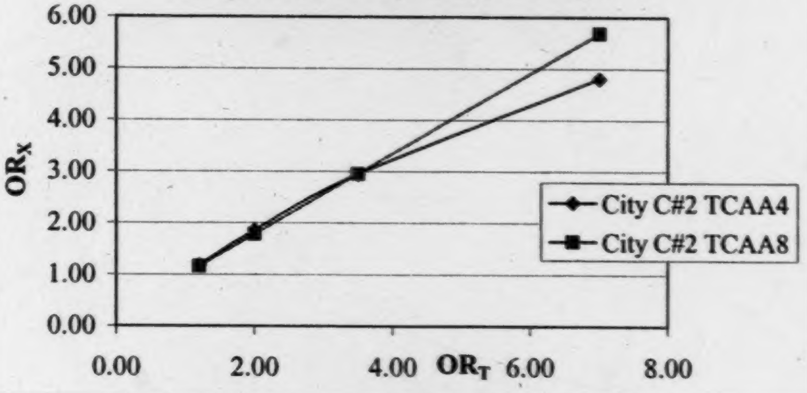


Figure 8-2. Example of the effect of sensitivity and specificity on OR_X

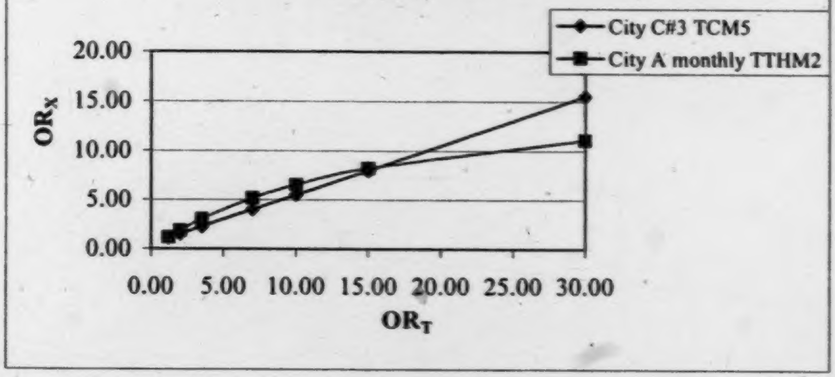
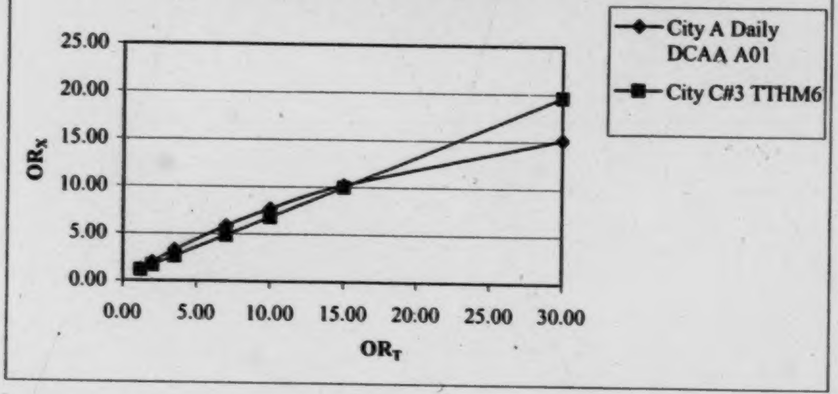


Fig. 8-3. Example of the effect of sensitivity and specificity on OR_X



Both Figures 8-2 and 8-3 show that there is a change-over point where the influence of sensitivity and specificity reverse. However, this change-over point is at highly unrealistic "true" ORs that are unlikely to happen in a "real life" situation. Therefore, with the exception of certain anomalies, the data cautiously support the observation by Reif et al. (2001) that specificity has a greater effect on OR attenuation than sensitivity.

We have discussed the generally greater OR attenuation seen in the categorical treatment over the continuous treatment. There are also cases where the OR attenuation in the categorical data is so great that the opposite relationship between exposure and effect can be inferred from the two data treatments. Examples are shown in Table 8-8. The examples given are from Chapter 5 only, although there are additional examples in Chapter 6.

None of the continuous r-values are very high. However, the last three examples would qualify as producing "measured" ORs within 10% of the "true" OR at OR_T=1.20. All the continuous results suggest a positive, if very minor, association between exposure and outcome. The measured ORs attenuated below one in the categorical treatment suggest a negative association between exposure and outcome. This is another example of the discrepancies that can occur between categorical and continuous treatment of exposure data.

Table 8-8. Summary of inconsistency between categorical and continuous treatment of the data

	Sensitivity	Specificity	Correlation coefficient (r)	OR_T= 1.20	OR_T= 2.00	OR_T= 3.50	OR_T= 7.00
Categorical							
City A daily ELS TCMA03	0.17	0.70		0.98	0.91	0.83	0.74
City A daily ELS DCANA03	0.00	0.65		0.95	0.78	0.59	0.38
City C#2 DCAA13	0.20	0.71		0.99	0.94	0.88	0.81
City C#3 BDCM7	0.00	0.80		0.95	0.81	0.63	0.41
Continuous							
City A daily ELS TCMA03			0.37	1.07	1.29	1.59	2.05
City A daily ELS DCANA03			0.56	1.11	1.47	2.01	2.95
City C#2 DCAA13			0.43	1.08	1.35	1.71	2.31
City C#3 BDCM7			0.66	1.13	1.58	2.29	3.63

Conclusion

The most important observation from this analysis is that categorical treatment can result in very different results compared to continuous treatment. In addition, the cutpoint chosen for categorical analysis has a large effect on the final ORs. There is widespread recognition that information is lost when continuous exposure measures are categorized, and continuous data should be used when possible. However, the practice of categorizing continuous data continues to be the norm in epidemiology studies. In fact, only two of the epidemiology studies (King et al., 2000, Savitz et al., 1995) reported results from continuous treatment of exposure data. Most of the epidemiology studies that have attempted to give an indication of a dose have divided continuous exposure measurements into several categories. In these studies, the categorical analyses were not dichotomous in that they included all the categories to give a measure of increasing dose. However, based on the large variations in the results from the different cutpoints in the dichotomous data, we would venture to say that treating the data categorically, even in a non-dichotomous analysis, could also result in widely varying conclusions. Therefore, if continuous data are available, the most believable results can be obtained from treating the data as continuous, rather than categorical. The argument for treating the data categorically has traditionally been that the data used (quarterly monitoring data from water utilities, for the most part) are subject to inaccuracies and do not adequately represent the individual exposure of the study subjects. Therefore, it was felt that concentration ranges or categories would best serve the purposes of the epidemiology studies. However, as extrapolated from the analysis in this chapter, the cutpoints chosen in categorical analysis can greatly influence the results of these studies.

The reasoning that dividing quarterly continuous monitoring data into categories will address the problem that the continuous data are not representative of individual exposure skirts the real issue of inadequate individual data. What is needed is the application of adequate continuous data, not inadequate categorical data. As discussed in Chapter 2 and noted in a recent weight-of-evidence analysis (Graves et al., 2001), DBP exposure assessment has not substantially improved in the decades since the possibility of an association between DBP exposure and adverse reproductive or developmental effects was first postulated.

The analyses done in this dissertation confirm that surrogate measures of sampling in individual homes and different DBP species are not ideal. However, they suggest that under certain circumstances surrogate measures can provide a better estimate of individual exposure than those measures employed currently. For example, water utility monitoring samples at the water treatment plants could be increased from quarterly to monthly. The analysis in Chapter 5 suggests that monthly water treatment plant DBP concentrations can provide a reasonable estimate of monthly distribution system DBP concentrations at lower ORs in both chlorinated and chloraminated systems, particularly for HAAs. Daily sampling at the water treatment plants would likely provide better information because monthly maximums and means could then be determined. However, the analysis in Chapter 7 concluded that there was good correlation between monthly means and maximums and convenience sampling. Monthly monitoring would benefit the water utilities as well, in giving a more complete picture of their water treatment processes.

Distribution system concentrations intuitively provide a better indication of individual DBP exposure by virtue of being spatially closer to the study subjects. However, there can be substantial differences in DBP concentrations between distribution sampling location and the individual homes of the study subjects, due to continually changing distribution system characteristics such as water demand, pump operation, and pipe diameter, among other factors. A proposed monitoring rule for DBP concentrations in the distribution system in the United States may address some of the variability in the distribution systems. The proposed new rule is based on a location running annual average (LRAA). This means that each location sampled must have a running annual average below the proposed concentration levels (ultimately 80μg/L for TTHM and 60 μg/L for HAA_5). The previous rule was based on a running annual average (RAA) for all sampling in the distribution system. This meant that all the DBP concentration variations in the distribution system were averaged out and some locations could conceivably exceed the regulated concentrations levels on a regular basis. The LRAA in the new proposed rule will ensure that locations with higher concentrations (such as those with a long residence time) will have to comply with the regulations. It is hoped that this will decrease spatial DBP concentration variation in the distribution system, although this remains to be seen. A decrease in DBP concentrations variations throughout a distribution system may aid in exposure assessment for future epidemiology studies.

In addition to the elements discussed above, findings in Chapters 5, 6, and 7 suggest adjustments to certain elements of exposure assessment that can improve the likelihood that the data are representative of individual exposure. In providing more representative exposure assessment, we can reduce exposure misclassification and the resulting OR attenuation. The analysis in this chapter suggests that using representative, continuous individual exposure data will result in less OR attenuation and clearer associations than categorizing inadequate exposure data.

CONTINUOUS

Table 8-1. Correlations between **daily** DBP concentration data at Water treatment plant #1 and home sampling locations in City A, and their effects on the odds ratio

TCM

	Correlation (r)	True Odds Ratio			
	TCM at WTP1 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
TCM at site A01	0.67	1.13	1.59	2.31	3.66
TCM at site A02	0.77	1.15	1.70	2.62	4.46
TCM at site A03	0.32	1.06	1.24	1.49	1.85
TCM at site A04	0.62	1.12	1.54	2.17	3.34

DCAN

	Correlation (r)	True Odds Ratio			
	TCM at WTP1 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
DCAN at site A01	0.59	1.11	1.51	2.09	3.15
DCAN at site A02	0.82	1.16	1.77	2.80	4.94
DCAN at site A03	0.46	1.09	1.37	1.77	2.42
DCAN at site A04	0.77	1.15	1.70	2.61	4.45

DCAA

	Correlation (r)	True Odds Ratio			
	TCM at WTP1 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
DCAA at site A01	0.95	1.19	1.94	3.30	6.40
DCAA at site A02	0.91	1.18	1.88	3.13	5.90
DCAA at site A03	0.86	1.17	1.82	2.94	5.33
DCAA at site A04	0.91	1.18	1.88	3.12	5.86

TCAA

	Correlation (r)	True Odds Ratio			
	TCM at WTP1 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
TCAA at site A01	0.67	1.13	1.59	2.30	3.65
TCAA at site A02	0.62	1.12	1.53	2.17	3.33
TCAA at site A03	0.50	1.10	1.42	1.88	2.67
TCAA at site A04	0.72	1.14	1.64	2.46	4.04

OR_T = the assumed OR for the true data

OR_X = the calculated OR for the measured data = OR_T^r

WTP1 = Water treatment plant #1

Sites A01-A04 are home sampling locations

HALVES

Table 8-2. Sensitivity and specificity between **daily** DBP concentration data at Water treatment plant #1 and distribution system locations in City A, and their effects on the odds ratio

TCM

			True Odds Ratio			
	TCM at WTP1 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
TCM at site A01	0.67	0.67	1.04	1.20	1.45	1.82
TCM at site A02	0.79	0.77	1.08	1.37	1.86	2.73
TCM at site A03	0.57	0.54	1.01	1.06	1.12	1.21
TCM at site A04	0.53	0.50	1.00	1.02	1.03	1.05

DCAN

			True Odds Ratio			
	DCAN at WTP1 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
DCAN at site A01	0.80	0.83	1.10	1.46	2.06	3.16
DCAN at site A02	0.80	0.83	1.10	1.46	2.06	3.16
DCAN at site A03	0.67	0.67	1.04	1.20	1.45	1.82
DCAN at site A04	0.67	0.67	1.04	1.20	1.45	1.82

DCAN

			True Odds Ratio			
	DCAA at WTP1 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
DCAA at site A01	0.87	0.92	1.14	1.67	2.61	4.47
DCAA at site A02	0.87	0.92	1.14	1.67	2.61	4.47
DCAA at site A03	0.80	0.83	1.10	1.46	2.06	3.16
DCAA at site A04	0.93	1.00	1.20	1.97	3.36	6.35

TCAA

			True Odds Ratio			
	TCAA at WTP1 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
TCAA at site A01	0.73	0.75	1.07	1.31	1.70	2.36
TCAA at site A02	0.73	0.75	1.07	1.31	1.70	2.36
TCAA at site A03	0.67	0.67	1.04	1.20	1.45	1.82
TCAA at site A04	0.80	0.83	1.10	1.46	2.06	3.16

OR_T = the assumed OR for the true data
OR_X = the calculated OR for the measured data
WTP1 = Water treatment plant #1
Sites A01-A04 are home sampling locations
P_N = 0.2; P_D (OR_T=1.20)=0.23; P_D (OR_T=2.00)=0.33; P_D (OR_T=3.50)=0.0.47; P_D (OR_T=7.00)=0.64;

TERTILES

Table 8-3. Sensitivity and specificity between **daily** DBP concentration data at Water treatment plant #1 and distribution system locations in City A, and their effects on the odds ratio

TCM

			True Odds Ratio			
	TCM at WTP1 (Measured)		$OR_T=1.20$	$OR_T=2.00$	$OR_T=3.50$	$OR_T=7.00$
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
TCM at site A01	0.70	0.88	1.10	1.48	2.08	3.10
TCM at site A02	0.89	0.94	1.15	1.74	2.77	4.89
TCM at site A03	0.40	0.71	1.02	1.07	1.14	1.24
TCM at site A04	0.80	0.94	1.14	1.68	2.60	4.31

DCAN

			True Odds Ratio			
	DCAN at WTP1 (Measured)		$OR_T=1.20$	$OR_T=2.00$	$OR_T=3.50$	$OR_T=7.00$
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
DCAN at site A01	0.80	0.94	1.14	1.68	2.60	4.31
DCAN at site A02	0.60	0.82	1.07	1.31	1.68	2.26
DCAN at site A03	0.70	0.88	1.10	1.48	2.08	3.10
DCAN at site A04	0.70	0.88	1.10	1.48	2.08	3.10

DCAA

			True Odds Ratio			
	DCAA at WTP1 (Measured)		$OR_T=1.20$	$OR_T=2.00$	$OR_T=3.50$	$OR_T=7.00$
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
DCAA at site A01	0.90	1.00	1.19	1.95	3.30	6.11
DCAA at site A02	0.70	0.88	1.10	1.48	2.08	3.10
DCAA at site A03	0.60	0.82	1.07	1.31	1.68	2.26
DCAA at site A04	0.90	1.00	1.19	1.95	3.30	6.11

TCAA

			True Odds Ratio			
	TCAA at WTP1 (Measured)		$OR_T=1.20$	$OR_T=2.00$	$OR_T=3.50$	$OR_T=7.00$
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
TCAA at site A01	0.70	0.88	1.10	1.48	2.08	3.10
TCAA at site A02	0.70	0.88	1.10	1.48	2.08	3.10
TCAA at site A03	0.30	0.65	0.99	0.97	0.94	0.91
TCAA at site A04	0.70	0.88	1.10	1.48	2.08	3.10

OR_T = the assumed OR for the true data
OR_X = the calculated OR for the measured data
WTP1 = Water treatment plant #1
Sites A01-A04 are home sampling locations
$P_N = 0.2$; P_D (OR_T=1.20)=0.23; P_D (OR_T=2.00)=0.33; P_D (OR_T=3.50)=0.0.47; P_D (OR_T=7.00)=0.64;

QUARTILES

Table 8-4. Sensitivity and specificity between **daily** DBP concentration data at Water treatment plant #1 and distribution system locations in City A, and their effects on the odds ratio

TCM

			True Odds Ratio			
	TCM at WTP1 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
TCM at site A01	0.71	0.90	1.11	1.53	2.19	3.34
TCM at site A02	0.71	0.90	1.11	1.53	2.19	3.34
TCM at site A03	0.29	0.75	1.01	1.03	1.06	1.09
TCM at site A04	0.71	0.90	1.11	1.53	2.19	3.34

DCAN

			True Odds Ratio			
	DCAN at WTP1 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
DCAN at site A01	0.57	0.85	1.07	1.34	1.73	2.34
DCAN at site A02	0.71	0.90	1.11	1.53	2.19	3.34
DCAN at site A03	0.27	0.75	1.00	1.01	1.03	1.05
DCAN at site A04	0.83	0.95	1.15	1.73	2.72	4.64

DCAA

			True Odds Ratio			
	DCAA at WTP1 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
DCAA at site A01	1.00	1.00	1.20	2.00	3.50	7.00
DCAA at site A02	0.71	0.90	1.11	1.53	2.19	3.34
DCAA at site A03	0.83	0.95	1.15	1.73	2.72	4.64
DCAA at site A04	0.71	0.90	1.11	1.53	2.19	3.34

TCAA

			True Odds Ratio			
	TCAA at WTP1 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
TCAA at site A01	0.83	0.95	1.15	1.73	2.72	4.64
TCAA at site A02	0.83	0.95	1.15	1.73	2.72	4.64
TCAA at site A03	0.27	0.75	1.00	1.01	1.03	1.05
TCAA at site A04	0.57	0.85	1.07	1.34	1.73	2.34

OR_T = the assumed OR for the true data
OR_X = the calculated OR for the measured data
WTP1 = Water treatment plant #1
Sites A01-A04 are home sampling locations
P_N = 0.2; P_D (OR_T=1.20)=0.23; P_D (OR_T=2.00)=0.33; P_D (OR_T=3.50)=0.0.47; P_D (OR_T=7.00)=0.64;

SEXTILES

Table 8-5. Sensitivity and specificity between daily DBP concentration data at Water treatment plant #1 and distribution system locations in City A, and their effects on the odds ratio

TCM

			True Odds Ratio			
	TCM at WTP1 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
TCM at site A01	0.60	0.95	1.13	1.60	2.32	3.50
TCM at site A02	0.60	0.95	1.13	1.60	2.32	3.50
TCM at site A03	0.00	0.82	0.96	0.81	0.63	0.42
TCM at site A04	0.60	0.95	1.13	1.60	2.32	3.50

DCAN

			True Odds Ratio			
	DCAN at WTP1 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
DCAN at site A01	0.40	0.91	1.07	1.34	1.71	2.25
DCAN at site A02	0.60	0.95	1.13	1.60	2.32	3.50
DCAN at site A03	0.20	0.86	1.01	1.06	1.13	1.21
DCAN at site A04	0.60	0.95	1.13	1.60	2.32	3.50

DCAA

			True Odds Ratio			
	DCAA at WTP1 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
DCAA at site A01	0.60	0.95	1.13	1.60	2.32	3.50
DCAA at site A02	0.80	1.00	1.19	1.91	3.13	5.44
DCAA at site A03	0.80	1.00	1.19	1.91	3.13	5.44
DCAA at site A04	0.80	1.00	1.19	1.91	3.13	5.44

TCAA

			True Odds Ratio			
	TCAA at WTP1 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
TCAA at site A01	0.40	0.91	1.07	1.34	1.71	2.25
TCAA at site A02	0.60	0.95	1.13	1.60	2.32	3.50
TCAA at site A03	0.20	0.86	1.01	1.06	1.13	1.21
TCAA at site A04	0.40	0.91	1.07	1.34	1.71	2.25

OR_T = the assumed OR for the true data
OR_X = the calculated OR for the measured data
WTP1 = Water treatment plant #1
Sites A01-A04 are home sampling locations
P_N = 0.2; P_D (OR_T=1.20)=0.23; P_D (OR_T=2.00)=0.33; P_D (OR_T=3.50)=0.0.47; P_D (OR_T=7.00)=0.64;

References

Graves, Carol G., Genevieve M. Matanoski and Robert G. Tardiff, 2001. Weight of Evidence for an Association between Adverse Reproductive and Evelopmental Effects and Exposure to Disinfection By-products: A Critical Review. *Regulatory Toxicology and Pharmacology* **34**: 103-124.

King, Will D., Linda Dodds and Alexander C. Allen, 2000. Relation Between Stillbirth and Specific Chlorination By-Products in Public Water Supplies. *Environmental Health Perspectives* **108**(9): 883-886.

Reif, John S., Annette Bachand and Melvin Andersen, 2000. *Reproductive and Developmental Effects of Disinfection By-Products*. Ottawa, Ontario. Health Canada. October 31, 2000.

Savitz, David A., Kurtis W. Andrews and Lisa M. Pastore, 1995. Drinking Water and Pregnancy Outcome in Central North Carolina: Source, Amount, and Trihalomethane Levels. *Environmental Health Perspectives* **103**(6): 592-596.

Chapter 9
Summary and Conclusions

In Chapter 1, a discussion was presented outlining the challenges of developing accurate exposure assessment for epidemiology studies investigating a possible causal association between exposure to disinfection byproducts in treated drinking water and adverse reproductive and developmental effects. Three major factors were identified as contributing to the difficulties in establishing accurate exposure assessment. The first of these factors is the formation of DBPs. The factors affecting the formation of disinfection by-products (DBPs), e.g. formation kinetics, disinfectant concentrations, available organic matter, contact time, temperature, and others, lead to variations in DBP concentrations over distance from the point of chlorination to the point of exposure and over time. Epidemiology studies that have considered DBP concentrations in their exposure assessment have generally utilized water treatment plant data or monthly or quarterly monitoring data that do not account for the potential for temporal and spatial variations in DBP concentrations. The second factor is the identification of the causal agent. While most epidemiology studies that have used DBP concentrations in their exposure assessment have used TCM or TTHM concentrations, toxicology studies point to other DBP species, such as HAAs, as more likely causal agents. The third factor relates to pathways of exposure to DBPs. Exposure to DBPs in treated drinking water is possible via ingestion, inhalation, and dermal contact. K_{ow} is an important factor in the determination of dermal intake; however, the K_{ow}s for many DBPs have not been experimentally determined. As part of the experimental work for this dissertation, log K_{ow} values for DCAN and CH were experimentally determined to be -0.59 and 1.46, respectively.

Of those epidemiology studies that have considered intake of DBPs in their exposure assessment, only the ingestion route was considered in final calculations. It seems obvious that exposure assessment can and should be improved. But, how good does exposure assessment have to be for these epidemiology studies? This question was postulated at recent international workshops on the issue of DBPs and adverse reproductive and developmental effects. The attempt to elucidate this question, at least in part, is at the core of the analyses and discussions in the subsequent chapters of this dissertation.

In Chapter 2, a framework was developed to group past and future studies according to their exposure assessment. Three basic aspects of epidemiological exposure assessment were identified and incorporated into the framework. The first aspect is personal vs. aggregated population information. This aspect of exposure assessment pertains to information on water supply, water consumption, water use, and other exposures. Personal information is obtained from interviews or questionnaires and can include questions regarding the water supply, water consumption, water use and other exposures. Aggregated population data links a study subject's residence to a geographic area served by a particular water utility or type of water (i.e. surface or ground) and generally provides information only on water supply. Personal information is considered to be more accurate in that it provides information at the individual level rather than a population level. The second aspect is the issue of using specific vs. non-specific chemicals. Data on specific agents includes the measurement of concentrations of individual compounds or groups of compounds. Information on non-specific agents includes the type and variability of water source or disinfection method. The measurement of specific agents is considered to be more accurate. The measurement of concentrations of specific chemicals allows the generation of a dose-response association. The third aspect is direct vs. indirect measurement of specific compounds. Direct measurement is spatially and temporally proximate to the individual study subjects. In other words, the DBP concentrations are measured at or close to the point of exposure and during the critical exposure period. An example of indirect measurement is quarterly monitoring data averaged over three monthly samples taken at the water treatment plant.

These three aspects were combined to define five exposure categories that are the basis of the exposure framework. All epidemiology studies done to date investigating a possible causal link between exposure to DBPs and adverse reproductive effects were assigned to one of the five categories based on their exposure assessment. The categories are arranged based on level of exposure assessment with Category 1 being the lowest level of exposure assessment. Category 5 is the highest level of exposure assessment, represented by biomarkers of exposure. Of the 23 studies conducted to date, seven were assigned to Category 1, seven to Category 2a, six to Category 2b, two to Category 3, and one to Category 4. There were no studies assigned to Category 5 because appropriate biomarkers of exposure have yet to be validated and utilized in epidemiology studies. Use of this framework shows that very few studies have employed the highest levels of exposure assessment incorporating both specific agents and personal data. Additionally, there is no discernable improvement in exposure assessment over time with the most recent studies classified in the lower exposure categories.

Several weaknesses of the studies as a body of work were also identified. There is a lack of consistent exposure levels and reference levels among studies looking at the same exposure elements. This issue should be addressed in future studies. The framework analysis suggests that the noise associated with less sophisticated methods of exposure assessment can overcome the minor improvements in exposure assessment between Categories 1, 2a, and 2b. Therefore, it is even more important for future studies to employ higher levels of exposure assessment. There is little merit in having more studies performed with inadequate exposure assessment. Future studies should plan their exposure assessments in light of the evidence available from prior studies, including the measures used for individual exposure assessment. It is recommended that future studies concentrate on exposure assessment employing specific agents and personal data on water consumption and water use activities.

Chapter 3 presents a discussion of exposure misclassification and its implications for epidemiology studies. Misclassification is discussed in terms of both continuous and categorical treatment of exposure data. Methods of quantifying the effect of exposure misclassification on the odds ratio (OR) of epidemiology studies are discussed. Measures of exposure misclassification between two data sets such as correlation coefficients for continuous data and sensitivity and specificity for categorical data are mathematically linked to OR attenuation. The equations discussed in Chapter 3 are the basis of the analysis and discussion of Chapters 5 through 7.

Chapter 4 is a discussion of the data sources and data collection methods used in this project. Daily DBP concentrations at the water treatment plants and in homes in the distribution systems over a month-long period in both City A and City B were determined experimentally during a sampling and analysis program developed and executed by the author. Monthly monitoring data was supplied by the water utilities in City A, City B, and City C. The experimental protocol for the K_{ow} determination is outlined in this Chapter, as are the results of these analyses. The sampling and analysis protocols for the daily sampling programs are also describe in this Chapter. In addition, the characteristics of the three water systems that are applicable to the analysis of these data are discussed.

Chapters 5, 6, and 7 present different ways of looking at the question "How good does exposure assessment have to be?" In all three chapters the mathematical methods discussed in Chapter 3 were used to illustrate the potential quantitative effect on exposure misclassification measures to OR attenuation.

Chapter 5 investigates the effect of spatial variations in DBP concentrations, particularly those between the water treatment plants and the distribution system.

Conclusions from Chapter 5:

- Data in a chloraminated system generally result in less exposure misclassification and less OR attenuation than did data from a chlorinated system.
- HAAs result in less exposure misclassification and subsequently less OR attenuation than do THMs or DCAN.
- Individual compounds (i.e. TCM) tend to result in less exposure misclassification and less OR attenuation than do groups of compounds (i.e. TTHM).

Ideally, detailed and specific exposure assessments for each study subject would form the basis of the exposure measure in epidemiology studies looking at a possible causal association between exposure to DBPs in drinking water and adverse reproductive effects. However, resource and feasibility constraints ensure that this will not be achieved for a full-scale field study. Therefore, from the general observations taken from these analyses of the data available for this study, we can make a general statement aimed towards minimizing exposure misclassification for these studies. If water treatment plant DBP concentrations are used as surrogate measures for distribution system DBP concentrations, then the least misclassification and least resulting OR attenuation is likely to occur if *HAA data* from a *chloraminated system* are used. As discussed in Chapter 5, there are exceptions to this general statement. However, if non-HAA compounds are measured in systems using chlorination, the data support current thinking that exposure misclassification due to spatial variations in DBP concentrations is best avoided by sample collection as close to the individual study subjects as possible.

Chapter 6 investigates the effect of measuring different causal agents. Most studies to date have used TCM or TTHM concentrations. Chapter 6 looks at the potential for exposure misclassification in using TCM or TTHM concentrations rather than toxicologically more plausible compounds, such as HAAs.

Conclusions from Chapter 6:

- For most conditions, TCM and TTHM result in less exposure misclassification and OR attenuation when compared with non-HAA data (other THMs, DCAN) than when compared with HAA data. A notable exception to this conclusion is the case of the monthly data from the chlorinated system.
- TCM and TTHM provide adequate surrogate measures for most other DBP species at the lowest OR_T (OR_T=1.20). However, this conclusion does not hold for higher OR_Ts.
- These two trends were clearer in the chloraminated system than the chlorinated system.

At low OR_Ts, measuring TCM or TTHM as a surrogate for other more toxicologically plausible causal agents will likely not result in substantial exposure misclassification and subsequent attenuated ORs, particularly in chloraminated systems over chlorinated systems. However, as the OR_T increases, the degree of attenuation also increases and the need to measure the actual causal agent rather than a surrogate becomes clear. This conclusion is more critical to chlorinated systems than chloraminated systems, particularly for daily sampling.

Chapter 7 investigates the effect of monthly sampling on exposure misclassification. Many water utilities employ monthly convenience sampling in their monitoring programs. These data are then used in epidemiology studies. However, it is unclear if convenience sampling can account for the DBP concentration variations over the month-long period. Chapter 7 endeavours to investigate this question. In addition, if daily sampling is employed, monthly mean values may be used in monitoring programs or as the exposure assessment in epidemiology studies. The potential for exposure misclassification due to the use of convenience sampling or monthly mean DBP concentrations is investigated.

Conclusions from Chapter 7:

- Results were consistent between the convenience sampling data, which suggests that the day chosen for convenience sampling is irrelevant.

Discrepancies between the continuous and categorical treatments were very evident in this chapter:

- The continuous data suggested that mean values provide useful surrogate measures for maximum and minimum values at all OR_Ts investigated here. Convenience sampling data provide useful surrogate measures for mean, maximum, and minimum values at the lower OR_Ts ($OR_T = 1.20$ and 2.00).
- The categorical data suggest that convenience sampling data and mean values are good surrogate measures only at the lowest OR_T ($OR_T=1.20$).

The decision to use monthly convenience sampling or monthly mean values as the exposure measure will depend on the type of data treatment and the expected OR_T.

Several conclusions were common to Chapters 5, 6, and 7:

- The quantitative extent of exposure misclassification on OR attenuation increased as the OR_T increased. This effect was seen for both the continuous and categorical treatment. The proportional effect of exposure misclassification depends in part on the expected "true" OR. This was demonstrated in Figures 5-1 and 5-2 and is expected from the OR attenuation equations.
- The extent of exposure misclassification and subsequent OR attenuation will depend on whether the data are treated as categorical or continuous. In several instances, there were substantial discrepancies between the categorical and continuous treatment of the same data set pairs.

The discussion in Chapter 8 addressed some of the discrepancies between the categorical and continuous data. A trend was seen in Chapters 5, 6, and 7 that the categorical (quartiles) analysis showed consistently greater OR attenuation than the continuous analysis. In addition, some categorical data set pairs showed such extreme OR attenuation that a negative association between exposure and outcome would be inferred rather than the positive association inferred for the continuous treatment of the same data set pairs. A sub-section of the data from Chapter 5 was divided into tertiles and sextiles and re-analyzed categorically. The results showed widely varying OR attenuation with no clear patterns with respect to sensitivity, specificity, or OR attenuation between cutpoints. The categorical ORs in the new cutpoints (as with the quartiles) were generally lower than those in the continuous treatment. These findings suggest that categorizing DBP concentration data in epidemiology studies can contribute to OR attenuation and as a result can cloud the final determination of association.

Most of the epidemiology studies that have tried to give a measure of dose have divided exposure into several categories. In these studies, the categorical analyses were not dichotomous and all the categories were included to give a measure of increasing dose. However, the large variations in the results from the different cutpoints in the dichotomous data suggests that treating the data categorically, even in a non-dichotomous analysis, could also result in widely varying conclusions, particularly when compared to continuous data treatment. Therefore, if continuous data are available, they should be used as continuous data rather than being divided into categories.

The argument for categorical data treatment has traditionally been that the data used (mostly quarterly monitoring data from water utilities) are subject to inaccuracies and do not adequately represent the individual exposure of the study subjects. It was felt that concentration ranges or categories would smooth out the effect of the inaccuracies in the continuous data. However, the results in Chapters 5 through 8 suggest that categorizing continuous data can in fact contribute to exposure misclassification and OR attenuation.

The real issue is the application of adequate continuous data rather than inadequate categorical data. The findings in Chapters 5, 6, and 7 suggest adjustments to certain elements of exposure assessment that can improve the likelihood that the data are representative of individual exposure. In providing more representative exposure assessment, we can reduce exposure misclassification and the resulting OR attenuation.

In Chapter 2, we saw that few epidemiology studies to date employed the higher levels of exposure assessment and we suggested that future studies strive to measure specific agents and obtain personal data. In Chapters 5 through 7, we endeavored to describe the potential for exposure misclassification from a variety of factors and to quantify the effect of that exposure misclassification on the OR of a hypothetical epidemiology study. In Chapter 8 we considered the potential for exposure misclassification and OR attenuation introduced by categorizing continuous DBP concentration data. We concluded that the categorization of inadequate exposure data can not substitute for the use of adequate and representative continuous exposure data.

So how good does exposure assessment really have to be? Ultimately, we found that the degree to which exposure assessment must be taken depends on several key factors, including the DBP species measured, the type of disinfection, the frequency of sampling or monitoring, and the type of data treatment. It seems that in some cases, namely in the chloraminated system investigated here, exposure assessment of the type used in many epidemiology studies, e.g. monthly monitoring of THMs at the water treatment plant, is not predicted to result in exposure misclassification serious enough to have a substantial effect on the resulting OR, particularly at low predicted "true" ORs. However, the conclusions as a whole in Chapters 5 through 7, confirm that the current situation of measuring unknown causal agents at sampling points that are spatially and temporally removed from the individual study subjects is unsatisfactory. As expected, the conclusions suggest that the measurement of known causal agents close to the point of exposure, and during the critical exposure time period is preferable, particularly at higher expected true ORs. Therefore, the recommendations made in Chapter 2 stand.

While the conclusions in Chapters 5,6, and 7 hold for the water systems analyzed here, the fact remains that only three water systems were studied. The generalizability of these findings to a larger range of water systems has not been tested. Therefore, it is not recommended that these results be applied directly to other chlorinated or chloraminated systems. The analyses done here should be carried out on other systems to confirm or disprove the generalizability of these findings.

It must also be emphasized that the exposure measures termed "true" in this study are not the actual true exposures for epidemiology studies. It was necessary for the purposes of the statistical method to label the exposure measures "true" and "measured". The "true" exposure was the measure that more closely represented the likely exposure of a study subject, while the "measured" exposure was the measure that represented the exposure less closely. However, neither the "true" nor the "measured" exposures used here can be said to accurately measure the real exposure of study subjects. Although individual exposure monitoring and biomarkers may come closest to measuring "true" exposure, any practical exposure assessment will be an approximation of "true" exposure.

As mentioned in the introduction, there are many factors affecting exposure assessment. In this thesis only one aspect, DBP concentration variation, was examined due to practical constraints on the research. However, it would be a worthwhile exercise to examine the effects of other aspects of exposure assessment, such as dermal uptake, inhalation, and variability in ingestion of DBPs. The contribution to exposure assessment and OR attenuation will differ for each route of exposure depending on DBP species. For example inhalation and dermal exposures are relevant for volatile and non-polar DBPs like the THMs, but these other exposure routes are not important for non-volatile, polar DBPs like the haloacetic acids.

The contribution of DBP concentration variation on OR attenuation within the greater variability in all aspects of exposure assessment is not known. As such, the limited focus of the work on DBP concentration variation precludes finding a definitive answer to the question "How good does exposure assessment have to be?" However, it was not expected that a definitive answer would be found in this thesis. Rather, the goal was to employ a multidisciplinary approach to quantify the effect of one aspect of exposure assessment on the OR of a study and so to begin to identify some boundaries on the question. This is the first multidisciplinary attempt to examine quantitatively the effect of exposure misclassification on the measure of strength of association in epidemiology studies. As such, the results are preliminary and necessarily limited. However, they provide a starting point for future research and the confirmation that future research of this type is warranted.

In addition, the findings of this work allow some comment on the results of past epidemiology studies. It was found in this work that at lower true ORs (OR_T), there is limited OR attenuation as a result of exposure misclassification. In fact, at the lowest OR_T, misclassification must be so extreme that the true and measured exposures must have a correlation coefficient of less than 0.42 for the OR to be attenuated more than 10% off the OR_T. The epidemiology studies to date tend to have low ORs, with few exceptions. In particular, this applies to studies in the higher exposure categories discussed in Chapter 2. These studies employed exposure measures close enough to the true exposure that correlation coefficients between the true and measured exposure would likely be greater than 0.42. As mentioned previously, at those r-values there is limited OR attenuation at low OR_Ts. It seems reasonable, then, to argue that the true ORs in these epidemiology studies are likely fairly low. Consequently, because of the limited OR attenuation at low OR_Ts, the measure ORs found in the studies done to date are likely close to the true ORs. In other words, provided that exposure assessments done to date are even moderately accurate, it appears unlikely that the low ORs observed are attenuated substantially from the true ORs that were being estimated.

Future epidemiology studies employing improved exposure assessment are still required. However, Chapter 5 discusses the impracticality of really good exposure assessment, in the absence of biomarkers of exposure, due to financial and resource constraints and burden on the study subjects. The statistical method as used in this thesis provides a potential approach to this problem. The findings of this work suggest that, under certain circumstances, surrogate measures of exposure can provide a reasonable estimate of exposure for epidemiology studies. These surrogate measures result in little OR attenuation from exposure misclassification because they rank the study subjects the same or similarly with respect to one another as the true exposure measure. This work has shown that surrogate measures of exposure can have minimal impact on OR attenuation; however, the circumstances identified in this work cannot be assumed to apply to all water systems.

The statistical method is simple to apply to any exposure data where there are matched pairs of a "true" exposure and a proposed surrogate measure. Therefore, a pilot study could easily be done before embarking on future epidemiology studies to determine the most effective surrogate measures of exposure that should be used. If there is ample historical water utility data, then the pilot study would be a simple mathematical exercise comparing, for example, water treatment plant data with distribution system data. However, more effective would be a study comparing, for example, water treatment plant DBP concentrations with a measure of ingested dose based on home DBP concentrations and consumption diaries. The ingested dose is the desired "true" measure, but it is difficult to obtain for a large study population, while the water treatment plant data is easily obtained. This method would provide a good basis of comparison for whether water treatment plant data could be used as a surrogate measure of the ingestion dose.

The combinations of "true" and "measured" data set pairs are endless and limited only by the resources available for pilot studies. However, a confirmed surrogate measure for the water system to be used in the epidemiology study can reduce the resources needed for exposure assessment in the epidemiology study. Therefore, it is recommended that approach outlined in this thesis be used to perform pilot studies on exposure misclassification before embarking on future epidemiology studies. Appropriate effort at the beginning can save time and resources further on in the epidemiology study. In addition, a surrogate measure with a good correlation to a "true" exposure increases the certainty of the exposure assessment and, ultimately, the credibility of the study results.

Suggestions for future research and epidemiology studies include:

- Exposure assessment at or as close as possible to the point of exposure, e.g. at the individual homes of study subjects, or at points in the distribution system close to the homes of study subjects. Some promising ongoing research includes the use of DBP concentration modeling in the distribution system.
- Daily monitoring at water treatment plants should be considered. This is already in progress at some water utilities and is particularly feasible to set up for THMs. Daily sampling would allow the determination of mean, maximum, and minimum monthly concentrations. These data would prove useful in future epidemiology studies as well as providing an ongoing indication of water treatment plant performance. At the very least, monthly convenience sampling should be considered to track variations in DBP concentrations over time more comprehensively than quarterly monitoring.
- Better measures of dose need to be developed than those currently in use. This will entail developing a better understanding of DBP concentrations at the point of exposure and research into the contribution of exposure routes other than ingestion. Ideally, one or several reliable biomarkers of exposure will be developed. However, until these biomarkers are validated, conventional methods are the most likely candidates for providing reasonable estimates of exposure to DBPs.
- Future studies should carefully weigh the effects of using categorical rather than continuous data. The results in Chapter 8 show that the two methods of data treatment can give widely varying results. In addition, categorizing DBP concentration data in an effort to compensate for inherent inaccuracies in the data skirts the real issue of inadequate exposure assessment. Therefore, rather than compounding the potential for exposure misclassification and OR attenuation by

categorizing the already continuous, but inaccurate, DBP concentration data, continuous treatment of data representative of individual exposure should become the norm.

- The statistical method used here should be applied to other water systems to determine the generalizability of the results.
- This statistical method should be used in the preliminary stages of future epidemiology studies to identify potential surrogate measures of exposure, in order to conserve resources in the epidemiology study and to increase the credibility of the study results.

In an ideal world, comprehensive individual exposure assessment would be economically feasible and easy to do, providing little burden on the study participant or the researchers. However, we do not live in an ideal world and must therefore work with the tools we have at hand. Although exposure assessment in the absence of biomarkers of exposure will never be "ideal", there is still much room for improvement. In terms of working with the tools at hand, there are several tools waiting to be picked up and others which we can develop with a little work.

Bibliography

Alvanja, M., I. Godstein and Susser M., 1978. A Case-Control Study of Gastrointestinal and Urinary Tract Cancer Mortality and Drinking Water Chlorination. in *Water Chlorination: Environmental Impact and Health Effects.* R. L. Jolley, H. Gorchev and D. H. Hamilton. Ann Arbor (MI), Ann Armor Science Publishers: 395-409.

Armstrong, Ben G., 1998. Effect of Measurement Error on Epidemiological Studies of Environmental and Occupational Exposures. *Occupational and Environmental Medicine* **55**: 651-656.

Armstrong, Bruce K., Emily White and Rodolfo Saracci, 1992. *Principles of Exposure Measurement in Epidemiology*. Oxford, Oxford University Press.

Aschengrau, A., S. Zierler and A. Cohen, 1993. Quality of Community Drinking Water and the Occurrence of Late Adverse Pregnancy Outcomes. *Archives of Environmental Health* **48**(2): 105-113.

Bader, Erin, Kenneth L. Froese and Steve E. Hrudey, Submitted. Determination of Elimination Half-Life for TCAA in a Controlled Human Exposure Study. .

Balko, Justin, Kenneth L. Froese and Steve E. Hrudey, 2001. *Is Drinking Coffee Worse Than We Think? Production of HAAs During Beverage Preparation*. Microbial/Disinfection By-Products Health Effects Symposium, Lisle, Illinois, American Water Works Association.

Beaglehole, R., R. Bonita and T. Kjellstrom, 1993. *Basic Epidemiology*. Geneva, World Health Organization.

Benoit, Frank M., Guy L. LeBel and Brian Jay, 2000. Evaluation of Techniques for the Removal of Chlorinated Disinfection By-Products form Drinking Water at the Point of Use. *Presented as a poster at An International Workshop on Exposure Assessment for Disinfection By-Products in Epidemiology Studies. Ottawa, Canada. May 7-10, 2000.*

Bove, F.J., M.C. Fulcomer, J.B. Koltz, J. Esmart, E.M. Dufficy and R.T. Zagraniski, 1992. *Report on Phase IV-A: Public Drinking Water Contamination and Birthweight Fetal Deaths, and Birth Defects, A Cross Sectional Study.* New Jersey Department of Health. 1992.

Bove, Frank J., Mark C. Fulcome, Judith B. Klotz, Jorge Esmart, Ellen M Dufficy and Jonathan E. Savrin, 1995. Public Drinking Water Contamination and Birth Outcomes. *American Journal of Epidemiology* **141**(9): 850-862.

Brenniman, G.R., J. Vasilomanolakis-Lagos, J. Amsel, N. Tsukasa and A.H. Wolff, 1980. Case-Control Study of Cancer Seaths in Illinois Sommunities Served by Chlorinated or Non-Chlorinated Water. in *Water Chlorination: Environmental Impact and Health Effects*. R. L. Jolley, W. A. Brungs and R. B. Cumming. Ann Arbor (MI), Ann Arbor Science Publishers: 1043-1057.

Brooke, D.N., A.J. Dobbs and N. Williams, 1986. Octanol: Water Partition Coefficients (P): Measurement, Estimation, and Interpretation, Particularly for Chemicals with $P>10^5$. *Ecotoxicology and Environmental Safety* **11**: 251-260.

Bunge, Annette L. and James N. McDougal, 1999. Dermal Uptake. in *Exposure to Contaminants in Drinking Water: Estimating Uptake through the Skin and by Inhalation*. S. S. Olin. Washington, D.C., CRC Press/ International Life Sciences Institute Press.

Cantor, K.P., R. Hoover, Harge P., Mason T.J., Silverman D.T. and Altman R., 1987. Bladder Cancer, Drinking Water Source andTap Water Consumption: a Case-Control Study. *Journal of the National Cancer Institute* **79**: 1269-1279.

Christ, Suzanne A., Eleanor J. Read, Judy A. Stober and M. Kate Smith, 1996. Developmental Effects of Trichloroacetonitrile Administered in Corn Oil to Pregnant Long-Evans Rats. *Journal of Toxicology and Environmental Health* **47**: 233-247.

Cicmanec, J.L., L.W. Condie, G.R. Olson and Shin-Ru Wang, 1991. 90-Day Toxicity Study of Dichloroacetate in Dogs. *Fundamental and applied Toxicology* **17**: 376-389.

Crebelli, Riccardo and Angelo Carere, 1989. Genetic Toxicology of 1,1,2-trichloroethylene. *Mutation Research* **221**: 11-37.

de Klerk, Nicholas H., Dallas R. English and Bruce K. Armstrong, 1989. A Review of the Effects of Random Measurement Error on Relative Risk Estimates in Epidemiological Studies. *International Journal of Epidemiology* **18**(3): 705-712.

Deane, Margaret, Shanna H. Swan, John A. Harris, David M. Epstein and Raymond R. Neutra, 1989. Adverse Pregnancy Outcomes in Relation to Water Contamination, Santa Clara County, California, 1980-1981. *American Journal of Epidemiology* **129**(5): 894-904.

Deane, Margaret, Shanna H. Swan, John A. Harris, David M. Epstein and Raymond R. Neutra, 1992. Adverse Pregnancy Outcomes in Relation to Water Consumption: A Re-analysis of Data from the Original Santa Clara County Study, California 1980-1981. *Epidemiology* **3**(2): 94-97.

Delpizzo, Vincent and Joanna L. Borghesi, 1995. Exposure Measurement Errors, Risk Estimate and Statistical Power in Case-Control Studies Using Dichotomous Analysis of a Continuous Exposure Variable. *International Journal of Epidemiology* **24**(4): 851-862.

Diehl, Alicia C., Gerald E. Speitel, James M. Symons, Stuart W. Krasner, Cordelia J. Hwang and Sylvia E. Barret, 2000. DBP Formation During Chloramination. *Journal of the American Water Works Association* **92**(6): 76-90.

Diehl, Alicia C., Gerard E. Jr. Speitel, James M. Sumons and Stuart W. Krasner, 1995. *Factors Affecting Disinfection By-Product Formation During Chloramination*. Water Quality Annual conference, Anaheim, California, American Water Works Association.

Dodds, Linda and Will King, 2001. Relation Between Trihalomethane Compounds and Birth Defects. *Occupational and Environmental Medicine* **58**: 443-446.

Dodds, Linda, Will King, Christy Woolcott and Jason Pole, 1999. Trihalomethanes in Public Water Supplies and Adverse Birth Outcomes. *Epidemiology* **10**(3): 233-237.

Dosemeci, Mustafa, Sholom Wacholder and Jay H. Lubin, 1990. Does Nondifferential Misclassification of Exposure Always Bias a True Effect Toward the Null Value? *American Journal of Epidemiology* **132**(4): 746-748.

Epstein, D.L., G.A. Nolen, J.L. Randall, S.A. Christ, E.J. Read, J.A. Stoberet al., 1992. Cardiopathic Effect of Dichloracetate in the Fetal Long-Evans Rat. *Teratology* **46**: 225-235.

Fantuzzi, Guglielmina, Elena Righi, Guerrino Predieri, Giorgia Ceppelli, Fabriziomaria Gobba and Gabriella Aggazzotti, 2001. Occupational Exposure to Trihalomethanes in Indoor Swimming Pools. *The Science of the Total Environment* **264**: 257-265.

Fenster, Laura, Gayle C. Windham, Shanna H. Swan, David M. Epstein and Raymond R. Neutra, 1992. Tap or Bottled Water Consumption and Spontaneous Abortion in a Case-Control Study of Reporting Consistency. *Epidemiology* **3**(2): 120-124.

Flegal, Katherine M., Penelope M. Keyl and F. Javier Nieto, 1991. Differential Misclassification Arising from Nondifferential Errors in Exposure Measurement. *American Journal of Epidemiology* **134**(10): 1233-1244.

Freedman, D.M., K.P. Cantor, Lee N.L., L.S. Chen, H.H. Lei and C.E. Ruhl, 1997. Bladder Cancer and Drinking Water: a Population-Based Case-Control Study in Washington County, Maryland. *Cancer Causes and Control* **8**: 738-744.

Froese, K.F., M.I. Sinclair, S.E. Hrudey 2002, In Press. Trichloroacetic Acid as a Biomarker of Human Exposure to DBPs in Drinking Water - Adelaide Trial. *Environmental Health Perspectives* **110**.

Gallagher, Michael D., John R. Nuckols, Lorann Stallones and David A. Savitz, 1998. Exposure to Trihalomethanes and Adverse Pregnancy Outcomes. *Epidemiology* **9**(5): 484-489.

Gordon, Sydney M., Lance A. Wallace, Patrick J. Callahan, Donald V. Kenny and Marielle C. Brinkman, 1998. Effect of Water Temperature on Dermal Exposure to Chloroform. *Environmental Health Perspectives* **106**(6): 337-345.

Gottleib, M.S., J.K. Carr and J.R. Clarkson, 1982. Drinking water and cancer in Louisiana: a retrospective mortality study. *American Journal of Epidemiology* **116**: 652-667.

Graves, Carol G., Genevieve M. Matanoski and Robert G. Tardiff, 2001. Weight of Evidence for an Association between Adverse Reproductive and Developmental Effects and Exposure to Disinfection By-products: A Critical Review. *Regulatory Toxicology and Pharmacology* **34**: 103-124.

Hennekens, Charles H. and Julie E. Buring, 1987. *Epidemiology in Medicine*. Boston, Little, Brown and company.

Hertz-Picciotto, Irva, Shanna H. Swan, Raymond R. Neutra and Steven J. Samuels, 1989. Spontaneous Abortions in Relation to Consumption of Tap Water: An Application of Methods from Survival Analysis to a Pregnancy Follow-Up Study. *American Journal of epidemiology* **130**(1): 79-93.

Hill, Austin Bradford, 1965. The Environment and Disease: Association or Causation? *Proceedings of the Royal Society of Medicine* **58**: 295-300.

Howard, Philip H., 1991. *Handbook of Environmental Fate and Exposure Data for Organic Chemicals*. Chelsea, Michigan, Lewis Publishers, Inc.

Howard, Philip H., 1997. *Handbook of Environmental Fate and Exposure Data for Organic Chemicals*. Chelsea, Michigan, Lewis Publishers, Inc.

Hrudey, S.E., W. Chen and C. Rousseaux, 1996. *Bioavailability*. New York, CRC-Lewis Publishers.

Jaakola, J.J.K., P. Magnus, A. Skrondal, B.-F. Hwang, G. Becher and E. Dybing, 2001. Foetal Growth and Duration of Gestation Relative to Water Chlorination. *Occupational and Environmental Medicine* **58**: 437-442.

Jacob, Sam, Abdel-Aziz H. Abdel-Aziz, Samia A. Shouman and Ahmed E. Ahmed, 1998. Effect of Glutathione Modulation on the Distribution and Transplacental Uptake of 2-{14C}-chloroacetonitrile (CAN) Quantitative Whole-Body Autoradiographic Study in Pregnant Mice. *Toxicology and Industrial Health* **14**(4): 533-546.

Jo, Wan K., Clifford P. Weisel and Paul J. Lioy, 1990. Routes of Chloroform Exposure and Body Burden from Showering with Chlorinated Tap Water. *Risk Analysis* **10**(4): 575-585.

Kallen, Bengt A.J. and Elisabeth Robert, 2000. Drinking Water Chlorination and Delivery Outcome--A Registry-Based Study in Sweden. *Reproductive Toxicology* **14**: 303-309.

Kanitz, Stefano , Yaurana Franco, Vittoriana Patrone, Marta Caltabellotta, Enrico Raffo, Cristina Riggi et al., 1996. Association Between Drinking Water Disinfection and Somatic Parameters at Birth. *Environmental Health Perspectives* **104**(5): 516-520.

Keegan, T., H. Whitaker, M.J. Nieuwenhuijsen, M.B. Toledano, P. Elliott, J. Fawell, M. Wilkinson, N. Best. 2001. Use of Routinely Collected Data on Trihalomethane in Drinking Water for Epidemiological Purposes. *Occupational and Environmental Medicine* **58**: 447-452.

Keen, Carl L., Janet Y. Uriu-Hare, Shanna H. Swan and Raymond R. Neutra, 1992. The Effects of Water Source on Reproductive Outcome in Fischer-344 Rats. *Epidemiology* **3**(2): 130-133.

Kelsey, Jennifer L., W. D. Thompson and Alfred S. Evans, 1986. *Methods in Observational Epidemiology*. New York, Oxford University Press.

Kim, Hekap and Clifford P. Weisel, 1998. Dermal Absorption of Dichloro-and Trichloroacetic Acids from Chlorinated Water. *Journal of Exposure Analysis and Environmental Epidemiology* **8**(4): 555-575.

King, Will D. and Loraine D. Marrett, 1996. Case-Control Study of Bladder Cancer and Chlorination By-Products in Treated Water (Ontario, Canada). *Cancer Causes and Control* **7**(6): 596-604.

King, Will D., Linda Dodds and Alexander C. Allen, 2000. Relation Between Stillbirth and Specific Chlorination By-Products in Public Water Supplies. *Environmental Health Perspectives* **108**(9): 883-886.

Klotz, Judith B. and Laurie A. Pyrch, 1999. Neural Tube Defects and Drinking Water Disinfection By-Products. *Epidemiology* **10**(4): 383-390.

Kramer, Michael D. , Charles F. Lynch, Peter Isacson and James W. Hanson, 1992. The Association of Waterborne Chloroform with Intrauterine Growth Retardation. *Epidemiology* **3**(4): 407-413.

Krasner, S. 1999. Chemistry of Disinfection By-Product Formation. in *Formation and control of Disinfection By-Products in Drinking Water*. P. C. Singer. Denver, Colorado, American Water Works Association: 424.

Krasner, Stuart W., 1999. *Chemistry and Occurrence of Disinfection By-Products*. The Safety of Water Disinfection: Balancing Chemical and Microbial Risks, Miami Beach, Florida, International Life Sciences Institute.

Krasner, Stuart W., Michael J. McGuire, Joseph G. Jacangelo, Nancy L. Patania, Kevin M. Reagan and E. Marco Aieta, 1989. The Occurrence of Disinfection By-Products in US Drinking Water. *Journal of the American Water Works Association* **81**(8): 41-53.

Lahl, Uwe, Mehmet Cetinkaya, Jurgen V. Duszeln, Bernd Gabel, Burkhard Stachel and Wolfram Thiemann, 1982. Health Risks for Infants Caused by Trihalomethane Generation During Chemical Disinfection of Feeding Utensils. *Ecology of Food and Nutrition* **12**: 7-17.

Land, Paul C., E.L. Owen and Harry W. Linde, 1981. Morphologic Changes in Mouse Spermatozoa after Exposure to Inhalational Anaesthetics during Early Spermatogenesis. *Anesthiology* **53**: 53-56.

Levesque, Benoit, Pierre Ayotte, Alain LeBlanc, Erid Dewailly, Denis Prud'Homme, Robert Lavoie et al., 1994. Evaluation of Dermal Respiratory Chloroform Exposure in Humans. *Environmental Health Perspectives* **102**(12): 1082-1087.

Lin, Tsair-Fuh and Shih-Wen Hoang, 2000. Inhalation Exposure to THMs from Drinking Water in South Taiwan. *The Science of the Total Environment* **246**: 41-49.

Linder, Ralph E., Gary R. Klinefelter, Lillian F. Strader, Juan D. Suarez and Cheryl J. Dyer, 1994. Acute Spermatogenic Effects of Bromoacetic Acids. *Fundamental and Applied Toxicology* **22**: 422-430.

Lynberg, Michele, J.R. Nuckols, Peter Langlois, David Ashley, Philip Singer, Pauline Mendola et al., 2001. Assessing Exposure to Disinfection By-products in Women of Reproductive Age Living in Corpus Christi, Texas, and Cobb County, Georgia: Descriptive Results and Methods. *Environmental Health Perspectives* **109**(6): 597-604.

Mackay, Donald, Wany Ying Shiu and Kuo Ching Ma, 1993. *Illustrated Handbook of Physical Chemical Properties and Environmental Fate for Organic Chemicals*. Boca Raton, Lewis Publishers.

Mackay, Donald, Wan-Ying Shiu and Kuo-Ching Ma, 1995. *Illustrated Handbook of Physical and chemical Properties and Environmental Fate for organic Chemicals*. New York, CRC Lewis Publishers.

Magnus, Per, Jouni J.K. Jaakkola, Anders Skrondal, Jan Alexander, Georg Becher, Truls Krogh et al., 1999. Water Chlorination and Birth Defects. *Epidemiology* **10**(5): 513-517.

McGeehin, Michael A. , John S. Reif, Judith C. Becher and Ellen J. Mangione, 1993. Case-Control Study of Bladder Cancer and Water Disinfection Methods in Colorado. *American Journal of Epidemiology* **138**(7): 492-501.

Mercer, Robert R., 1999. Respiratory Uptake. in *Exposure to Contaminants in Drinking Water: Estimating Uptake through the Skin and by Inhalation*. S. S. Olin. Washington, D.C., CRC Press/ International Life Sciences Institute Press.

Mills, Christina J., Richard J. Bull, Kenneth P. Cantor, John Reif, Steve E. Hrudey and Patricia Huston, 1998. Health Risks of Drinking Water Chlorination By-Products: Report of an Expert Working Group. *Chronic Diseases in Canada* **19**(3): 91-102.

Morris, Robert D., Anne-Marie Audet, Italo F. Angelillo, Thomas C. Chalmers and Frederick Mosteller, 1992. Chlorination, Chlorination By-Products, and Cancer: A Meta-Analysis. *American Journal of Public Health* **82**(7): 955-963.

Murray, F.J., B.A. Schwetz, J.G. McBride and R.E. Staples, 1979. Toxicity of Inhaled Chloroform in Pregnant Mice and Their Offspring. *Toxicology and Applied Pharmacology* **50**: 515-522.

Narotsky, Michael G., Rex A. Pegram and Robert J. Kavlock, 1997. Effect of Dosing Vehicle on the Developmental Toxicity of Bromodichloromethane and Carbon Tetrachloride in Rats. *Fundamental and Applied Toxicology* **40**: 30-36.

National Research Council, 1987. Biomarkers in Environmental Health Research. *Environmental Health Perspectives* **74**: 3-9.

National Research Council, 1989. *Biologic Markers in Reproductive Toxicology*. Washington, D.C., National Academy Press.

National Research Council, 1991. *Human Exposure Assessment for Airborne Pollutants: Advances and Opportunities*. Washington, D.C. National Research Council.

Nieuwenhuijsen, Mark J., Mireille B. Toledano and Paul Elliott, 2000. Uptake of Chlorination Disinfection By-Products; A Review and a Discussion of its Implications for Exposure Assessment in Epidemiological Studies. *Journal of Exposure Analysis and Environmental Epidemiology* **10**: 586-599.

Nieuwenhuijsen, Mark J., Mireille B. Toledano, Naomi E. Eaton, John Fawell and Paul Elliott, 2000a. Chlorination Disinfection By-Products in Water and Their Association with Adverse Reproductive Outcomes: A Review. *Occupational and Environmental Medicine* **57**: 73-85.

Paustenbach, Dennis J., 2000. The Practice of Exposure Assessment: A State-of-the-Art Review. *Journal of Toxicology and Environmental Health, Part B* **3**: 179-291.

Petrucci, Ralph H., 1989. *General Chemistry*. New York, Macmillan Publishing Company.

Reif, John S. , Maureen C. Hatch, Michael Bracken, Lewis B. Holmes, Bernard A. Schwetz and Philip C. Singer, 1996. Reproductive and Developmental Effects of Disinfection By-Products in Drinking Water. *Environmental Health Perspectives* **104**(10): 1056-1061.

Reif, John S., Annette Bachand and Melvin Andersen, 2000. *Reproductive and Developmental Effects of Disinfection By-Products*. Ottawa, Ontario. Health Canada. October 31, 2000.

Richardson, Susan D., 1998. Drinking Water Disinfection By-Products. in *Encyclopedia of Environmental Analysis and Remediation*. R. A. Meyers. New York, John Wiley & Sons: 1398-1421.

Rizak, Samantha, Steve E. Hrudey, Kelly Kjartanson, Trevor Satchwill and Kenneth L. Froese, 2000. *Distribution System Monitoring Data as a Basis to Estimate Human Exposure to Halogenated Disinfection By-products*. Ninth National Conference on Drinking Water, Ottawa, Ontario, Canadian Water and Wastewater Association.

Rook, J.J., 1974. Formation of Haloforms During Chlorination of Natural Waters. *Water Treatment and Examination* **23**: 234-243.

Rossman, Lewis A., Richard A. Brown, Philip C. Singer and John R. Nuckols, 1999. *Kinetics of THM and HAA Production in a Simulated Distribution System*. Water Quality Technology Conference, Tampa, Florida, American Water Works Association.

Rothman, Kenneth J. and Sander Greenland, 1998. *Modern Epidemiology*. Philadelphia, Lippincott-Raven Publishers.

Ruddick, J.A., D.C. Villeneuve and I. Chu, 1983. A Teratological Assessment of Four Trihalomethanes in the Rat. *Journal of Environmental Science and Health* **B18**(3): 333-349.

Savitz, David A., Kurtis W. Andrews and Lisa M. Pastore, 1995. Drinking Water and Pregnancy Outcome in Central North Carolina: Source, Amount, and Trihalomethane Levels. *Environmental Health Perspectives* **103**(6): 592-596.

Sawyer, Clair N., Perry L. McCarty and Gene F. Parkin, 1994. *Chemistry for Environmental Engineering*. New York, McGraw-Hill, Inc.

Schwetz, B.A., B.K.J. Leong and P.J. Gehring, 1974. Embryo- and Fetotoxicity of Inhaled Chloroform in Rats. *Toxicology and Applied Pharmacology* **28**: 442-451.

Shimokura, Gayle H., David A. Savitz and Elaine Symanski, 1998. Assessment of Water Use for Estimating Exposure to Tap Water Contaminants. *Environmental Health Perspectives* **106**(2): 55-59.

Singer, P.C. 1999. *Variability and Assessment of Disinfection By-Product Concentrations in Water Distribution Systems*. The Safety of Water Disinfection: Balancing Chemical and Microbial Risks, Miami Beach, Florida, International Life Sciences Institute.

Singer, P.C., A. Obolensky, A. Greiner. 1995. DBPs in Chlorinated North Carolina Drinking Waters. *Journal of the American Water Works Association* **87**(10): 83-92.

Smith, M.K., J.L. Randall, D.R. Tocco, R.G. York, J.A. Stober and E.J. Read, 1988. Teratogenic Effects of Trichloroacetonitrile in the Long-Evans Rat. *Teratology* **38**: 113-120.

Smith, M.K., J.L. Randall, E.J. Read and J.A. Stober, 1989a. Teratogenic Activity of Trichloroacetic Acid in the Rat. *Teratology* **40**: 445-451.

Smith, M.K., J.L. Randall, E.J. Read and J.A. Stober, 1992. Developmental Toxicity of Dichloroacetate in the Rat. *Teratology* **46**: 217-223.

Smith, M.K., J.L. Randall, J.A. Sotber and E.J. Read, 1989b. Developmental Toxicology of Dichloroacetonitrile: A By-Products of Drinking Water Disinfection. *Fundamental and Applied Toxicology* **12**: 765-772.

Snoeyink, Vernon L. and David Jenkins, 1980. *Water Chemistry*. New York, John Wiley and Sons, Inc.

Stevens, Alan A., Leown A. Moore and Richard J. Miltner, 1989. Formation and Control of Non-Trihalomethane Disinfection By-Products. *Journal of the American Water Works Association* **81**(8): 54-60.

Swan, Shanna H. and Kirsten Waller, 1998. Disinfection By-Products and Adverse Pregnancy Outcomes: What Is the Agent and How Should it Be Measured? *Epidemiology* **9**(5): 479-481.

Swan, Shanna H., Kirsten Waller, Barbara Hopkins, Gayle Windham, Laura Fenster, Catherine Schaefer et al., 1998. A Prospective Study of Spontaneous Abortion: Relation to Amount and Source of Drinking Water Consumed in Early Pregnancy. *Epidemiology* **9**(2): 126-133.

Thompson, D.J., S.D. Warner and V.B. Robinson, 1974. Teratology Studies on Orally Administered Chloroform in the Rat and Rabbit. *Toxicology and Applied Pharmacology* **29**: 348-357.

Toth, G.P., K.C. Kelty, E.L. George, E.J. Read and M.K. Smith, 1992. Adverse Male Reproductive Effects following Subchronic Exposure of Rats to Sodium Dichloroacetate. *Fundamental and Applied Toxicology* **19**: 57-63.

Tuthill, R.W., R.A. Giusti, G.S. Moore and E.J. Calabrese, 1982. Health Effects Among Newborns after Prenatal Exposure to ClO_2-Disinfected Drinking Water. *Environmental Health Perspectives* **46**: 39-45.

U.S. EPA, 1992a. *Dermal Exposure Assessment: Principles and Applications*. Washington, D.C. U.S. EPA, Office of Health and Environmental Assessment, Office of Research and Development. EPA/600/8-91/011.

U.S. EPA, 1992b. Guidelines for Exposure Assessment. *Federal Register* **57**(104): 22888-22938.

U.S. EPA, 1995a. *Method 551.1 Determination of Chlorination Disinfection Byproducts, Chlorinated Solvents, and Halogenated Pesticides/Herbicides in Drinking Water by Liquid-Liquid Extraction and Gas Chromatography with Electron Capture Detection.* Cincinnati, Ohio. US Environmental Protection Agency. 1995.

U.S. EPA, 1995b. *Method 552.2 Determination of Haloacetic Acids and Dalapon in Drinking Water by Liquid-Liquid Extraction, Derivatization and Gas Chromatography with Electron Capture Detection*. Cincinnati, Ohio. US Environmental Protection Agency. 1995.

U.S. EPA, 1997. *Exposure Factors Handbook, Volume I, General Factors.* Washington, D.C. U.S. EPA. 1997. EPA/600/P-95/002Fa.

U.S. Public Health Service, 1964. *Smoking and Health: Report of the Advisory Committee to the Surgeon General of the Public Health Service.* Washington, D.C. U.S. Public Health Service. PHS Publication No. 1103.

Valentine, Richard L., Brian Anferman, Shelley Hackett, Peter Vikesland and Skipp Slattenow, 1999. *Characterization of Disinfectant Decay and DBP Formation in the Presence of Water Distribution System Deposits.* Water Quality Technology Conference, Tampa, Florida, American Water Works Association.

van Haelst, A.G., P.F. Heese, F.W.M. van der Wielen and H.A.J. Govers, 1994. Determination of n-Octanol/Water Partition Coefficients of Tetrachlorobenzyltoluenes Individually and in a Mixture by the Slow Stirring Method. *Chemosphere* **29**(8): 1651-1660.

Wacholder, Sholom, Mustafa Dosmeci and Jay H. Lubin, 1991. Blind Assignment of Exposure Does Not Always Prevent Differential Misclassification. *American Journal of Epidemiology* **134**(4): 433-437.

Wallace, L.A., 1997. Human Exposure and Body Burden for Chloroform and Other Trihalomethanes. *Critical Reviews in Environmental Science and Technology* **22**(2): 113-194.

Wallace, Lance A., Edo Pellizzari, Ty Hartwell, Martin Rosenzweig, Mitchell Erickson, Charles Sparacino et al., 1984. Personal Exposure to Volatile Organic Compounds I. Direct Measurements in Breathing-Zone Air, Drinking Water, Food, and Exhaled Breath. *Environmental Research* **35**: 293-319.

Wallace, Lance, Edo Pellizzari, Ty Hartwell, Harvey Zelon, Charles Sparacino, Rebecca Perritt et al., 1986. Concentrations of 20 Volatile Organic Compounds in the Air and Drinking Water of 350 Residents of New Jersey Compared with Concentrations in their Exhaled Breath. *Journal of Occupational Medicine* **28**(8): 603-608.

Waller, Kirsten , Shanna H. Swan, Gerald DeLorenze and Barbara Hopkins, 1998. Trihalomethanes in Drinking Water and Spontaneous Abortion. *Epidemiology* **9**(2): 134-140.

Waller, Kirsten, Shanna H. Swan, Gayle C. Windham and Laura Fenster, 2001. Influence of Exposure Assessment Methods on Risk Estimates in an Epidemiologic Study of Total Trihalomethane Exposure and Spontaneous Abortion. *Journal of Exposure Analysis and Environmental Epidemiology* **11**: 522-531.

Weisel, Clifford and Wan-Kuen Jo, 1996. Ingestion, Inhalation, and Dermal Exposures to Chloroform and Trichloroethene from Tap Water. *Environmental Health Perspectives* **104**(1): 48-51.

Weisel, Clifford P., Hekap Kim, Patricia Haltmeier and Judith B. Klotz, 1999a. Exposure Estimates to Disinfection By-Products of Chlorinated Drinking Water. *Environmental Health Perspectives* **107**(2): 103-110.

Weisel, Clifford P., John C. Little, Nancy Chiu, Spyros N. Pandis, Cliff Davidson and Charles R. Wilkes, 1999b. Developing Exposure Estimates. in *Exposure to Contaminants in Drinking Water: Estimating Uptake through the Skin and by Inhalation*. S. S. Olin. Washington, D.C., CRC Press; International Life Sciences Institute.

Wilcosky, Timothy C., 1990. Criteria for Selecting and Evaluating Markers. in *Biological Markers in Epidemiology*. B. S. Huylka, T. C. Wilcosky and J. D. Griffith. New York, NY, Oxford University Press.

Wilkins, J.R. and G.W. Comstock, 1981. Source of Drinking Water at Home and Site-Specific Cancer Incidence in Washington County, Maryland. *American Journal of Epidemiology* **114**: 178-190.

Windham, Gayle C., Shanna H. Swan, Laura Fenster and Raymond R. Neutra, 1992. Tap or Bottled Water Consumption and Spontaneous Abortion: A 1986 Case-Control Study in California. *Epidemiology* **3**(2): 113-119.

World Health Organization, 1993. *Biomarkers and Risk Assessment: Concepts and Principles*. Finland, World Health Organization.

Wrensch, Margaret, Shanna H. Swan, Jane Lipscomb, David M. Epstein, Raymond R. Neutra and Laura Fenster, 1992. Spontaneous Abortions and Birth Defects Related to Tap and Bottled Water Use, San Jose, California 1980-1985. *Epidemiology* **3**(2): 98-103.

Wu, Marion L., Alice S. Whittemore and Dexter L. Jung, 1986. Errors in Reported Dietary Intakes. I. Short Term Recall. *American Journal of Epidemiology* **124**(5): 826-835.

Yang, Chun-Yuh, Bi-Hua Cheng, Shang-Shyue Tsai, Trong-Neng Wu, Meng-Ciao Lin and Kuo-Cheng Lin, 2000. Association Between Chlorination of Drinking Water and Adverse Pregnancy Outcome in Taiwan. *Environmental Health Perspectives* **108**(8).

Young, T.B., D.A. Wolf and M.S. Kanarek, 1981. Epidemiologic Study of Drinking Water Chlorination and Wisconsin Female Cancer Mortality. *Journal of the National Cancer Institute* **67**: 1191-1198.

Zierler, Sally, Lisa Feingold, Robert A. Danley and Gunther Craun, 1988. Bladder Cancer in Massachusetts Related to Chlorinated and Chloraminated Drinking Water: A Case-Control Study. *Archives of Environmental Health* 43(2): 194-200.

Appendix for Chapter 3

Logistic Regression

and

Tables 3-6 and 3-10

Logistic Regression Primer

Logistic regression is used to predict a *binary* or *categorical* variable ($y=1$ or $y=0$) from several independent variables ($x_1, x_2, x_3 \ldots$) that can be either *categorical* or *continuous*.

Logistic regression is used for this purpose because of the special properties of the function

$$\mathbf{f(z) = 1/(1+e^{-z})}$$

1) f(z) lies between 0 and 1, a requirement of a function describing a risk or probability
2) f(z) is S-shaped: f(z) is low for small values of z until a threshold is reached, after which the value of f(z) increases rapidly for intermediate values of z and remains around 1 for large values of z.

z represents the combination of **risk factors** used to predict the outcome of interest or,

$$z = \alpha + \Sigma\beta_i X_I = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 \ldots \beta_k X_k$$

Therefore,

$$\mathbf{f(z) = 1 / [1 + e^{-(\alpha+\Sigma\beta i X i)}]}$$

f(z) can also be written as a probability function,

$$f(z) = P(D=1/X_1, X_2, \ldots X_k)$$

Where $P(D=1/X_1, X_2, \ldots X_k)$ means *"the probability that D (the outcome) will be 1, given the factors $X_1, X_2, \ldots X_k$."*

This probability function can be written as:

$$\mathbf{P(D=1/X_1, X_2, \ldots X_k) = 1 / [1 + e^{-(\alpha+\Sigma\beta i X i)}]}$$

$P(D=1/X_1, X_2, \ldots X_k)$ can be abbreviated as $P(X_1)$

Computing Odds Ratios

This method is best explained by way of an example, i.e. the risk of getting a disease, say coronary heart disease (CHD), if you have high blood pressure (X_1) vs if you have normal blood pressure (X_0).

The odds of getting a disease if you have high blood pressure (X_1) can be calculated as *the probability of getting the disease, divided by the probability of not getting the disease*:

$$\text{Odds for } X_1 = P(X_1) / [1-(P(X_1)]$$

Or

$$= \{1 / [1 + e^{-(\alpha+\beta X1)}] \}* \{[e^{-(\alpha+\beta X1)}]/ [1 + e^{-(\alpha+\beta X1)}]\}$$

$$\textbf{Odds for } \mathbf{X_1} = e^{\alpha+\beta X1}$$

The odds for X_0 can be calculated in a similar fashion. The odds ratio (OR) will then be the odds of getting the disease if you have high blood pressure divided by the odds of getting the disease if you have normal blood pressure:

$$OR = e^{\alpha+\beta X1} / e^{\alpha+\beta X0}$$

$$\mathbf{OR} = e^{\beta(X1-X0)}$$

Applications with more than one variable

Most diseases have *more than one* contributing factor, therefore it is necessary to fit more terms to the model. This can be easily done by adding those variables and their β-coefficients to the equation.

For example, what is the risk of getting heart disease given the factors blood pressure (BP), age (AGE), and catecholine levels (CAT), where BP can be high (1) or low (0), CAT can be high (1) or low (0), and AGE is a continuous variable.

$$OR = e^{\alpha+\beta 1BP1+\beta 2AGE1+\beta 3CAT1} / e^{\alpha+\beta 1BP0+\beta 2AGE0+\beta 3CAT0}$$

$$\mathbf{OR} = e^{\beta 1(BP1-BP0)+\beta 2(AGE1-AGE0)+\beta 3(CAT1-CAT0)}$$

Or in general terms:

$$\mathbf{OR} = e^{\Sigma\beta i(x1i-X0i)}$$

The OR can now be calculated for BP, *controlling for* AGE and CAT, where BP = 1,0; AGE = 40 and CAT = 0:

$$OR = e^{\beta 1(1-0)+\beta 2(40-40)+\beta 3(0-0)}$$

$$\mathbf{OR} = e^{\beta 1}$$

Similarly, OR for one unit (year) increase in AGE, controlling for BP and CAT will be $e^{\beta 2}$, and OR for CAT controlling for BP and AGE will be $e^{\beta 3}$.

Confidence Intervals

Confidence intervals can be computed from the equation:

$$\mathbf{CI\ (95\%) = e^{\beta \pm 1.96SE}}$$

Where β is the regression coefficient and SE is the standard error, both calculated in the logistic regression model.

Logit Function

The logit function is the linear expression of logistic regression. The general formula for the logit function is:

$$\text{logit}(X) = \log[X/(1-X)]$$

Applying this concept to logistic regression:

$$\begin{aligned}\text{logit}[P(X)] &= \log\{P(X)/[1-P(X)]\}\\ &= \log(\text{odds for } X)\\ &= \alpha + \beta X\end{aligned}$$

Then we can say that:

$$\begin{aligned}\text{logit}[P(X)] - \text{logit}[P(X_0)] &= \log(\text{odds for } X) - \log(\text{odds for } X_0)\\ &= \log(OR_{X1,X0})\end{aligned}$$

$$\mathbf{logit\,[P(X)] - logit\,[P(X_0)] = \beta(X_1 - X_0)}$$

Four Main Equations in Logistic Regression

1) **$P(D=1/X_1) = 1 / [1 + e^{-(\alpha+\beta X)}]$**
2) **Odds for $X_1 = = e^{\alpha+\beta X1}$**
3) **Odds Ratio $= = e^{\beta(X1-X0)}$**
4) Logit $[P(X_1)] = \alpha + \beta X_1$

Special consideration for continuous variables

The OR calculated above for AGE is the increase in risk for *one* increment (i.e. one year) change in AGE. It is possible to make this measure more realistic by calculating the OR for a *larger* increment change, say c=10 years:

$$OR=e^{c\beta}$$

$$OR=e^{10\beta}$$

The confidence intervals for this OR are calculated as follows:

$$CI\ (95\%) = e^{c(\beta \pm 1.96SE)}$$

The assumption governing this OR is that the increase in risk is the same throughout the entire lifespan. This assumption can (and should) be tested, however, this test is not be covered in this summary.

Interaction and Confounding

Variables should be tested for interaction and confounding effects. Interaction should be tested first before confounding.

Interaction between two variables can be assessed using the ***Likelihood Ratio test*** or ***Wald test***. These tests use numbers generated by the logistic regression model. In the Likelihood Ratio test, it is important to fit two models. The first model will contain only the variables on their own. The second model will contain the variables alone and as a crossproduct of each other. The Likelihood Ratio test is usually preferred over the Wald test, particularly when more than one variable is being tested. Both tests are similar to a chi-squared statistic and its corresponding p-value.

Confounding between two variables can be tested by fitting two models and calculating the $\Delta\beta$. The first model contains both variables and gives us a regression coefficient, β_1. The second model contains only one variable and gives us a regression coefficient, β_2. These values are then used in the formula:

$$\Delta\beta = [(\beta_1-\beta_2)/\beta_1]*100$$

If $\Delta\beta > 20\%$, then there is a confounding effect between the two variables;

If $\Delta\beta < 20\%$, then there is no confounding effect between the two variables.

between the exposure and the disease.

Mathematical Proofs for Attenuation Equation

Proofs for variance and validity coefficient relationships (Allen and Yen, 1979)

To prove $\sigma^2_X = \sigma^2_T + \sigma^2_E$:

$\sigma^2_X = \sigma^2_{T+E}$ (From the assumption $X = T + E$)

$\sigma^2_X = \sigma^2_T + \sigma^2_E + 2\sigma_{ET}$

Because $\sigma_{ET} = 0$,

$\sigma^2_X = \sigma^2_T + \sigma^2_E$

To prove $\rho^2_{TX} = \sigma^2_T/\sigma^2_X$:

$\rho^2_{TX} = [\sigma_{XT}/(\sigma_X\sigma_T)]^2$

$\rho^2_{TX} = [\varepsilon(XT) - \varepsilon(X)\varepsilon(T)]^2/\sigma^2_X\sigma^2_T$

$\rho^2_{TX} = [\varepsilon\{(T + E)T\} - \varepsilon\,(X)\varepsilon(T)]^2/\sigma^2_X\sigma^2_T$

$\rho^2_{TX} = [\varepsilon(T2) + \varepsilon(ET) - \{\varepsilon(T)\}^2]^2/\sigma^2_X\sigma^2_T$

$\rho^2_{TX} = [\varepsilon(T^2) - \{\varepsilon(T)\}^2]^2/\sigma^2_X\sigma^2_T$

$\rho^2_{TX} = (\sigma^2_T)^2/\sigma^2_X\sigma^2_T$

$\rho^2_{TX} = \sigma^2_T/\sigma^2_X$

To prove $\rho^2_{XT} = 1 - (\sigma^2_E/\sigma^2_X)$:

$\rho^2_{XT} = \sigma^2_T/\sigma^2_X$

$\rho^2_{XT} = (\sigma^2_X - \sigma^2_E)/\sigma^2_X$

$\rho^2_{XT} = 1 - (\sigma^2_E/\sigma^2_X)$

To prove $\beta_X = [1 + (b_D - b_N)/(\mu_{TD} - \mu_{TN})]\ \beta_T\rho^2_{TX}$:

$\beta_X = [(\mu_{TD} - \mu_{TN}) + (b_D - b_N)] / (\sigma^2_T + \sigma^2_E)$

Substitute $\sigma^2_X = \sigma^2_T + \sigma^2_E$:

$\beta_X = [(\mu_{TD} - \mu_{TN}) + (b_D - b_N)]/\sigma^2_X$

Substitute $\rho^2_{TX} = \sigma^2_T/\sigma^2_X \Rightarrow \sigma^2_X = \sigma^2_T/\rho^2_{TX}$:

$\beta_X = [(\mu_{TD} - \mu_{TN}) + (b_D - b_N)]\rho2_{TX}/\sigma2_T$

$\beta_X = [(\mu_{TD} - \mu_{TN})\rho^2_{TX}/\sigma^2_T] + [(b_D - b_N)\rho^2_{TX}/\sigma^2_T]$

Multiply second term by $[(\mu_{TD} - \mu_{TN})/(\mu_{TD} - \mu_{TN})]$

and substitute $\beta_T = (\mu_{TD} - \mu_{TN})/\sigma^2_T$:

$\beta_X = \beta_T\rho^2_{TX} + \beta_T\rho^2_{TX}[(b_D - b_N)/(\mu_{TD} - \mu_{TN})]$

Finally:

$\beta_X = [1 + (b_D - b_N)/(\mu_{TD} - \mu_{TN})]\ \beta_T\rho^2_{TX}$

Table 3- 6. Effect of non-differential misclassification of categorical data

P_N=the proportion exposed in the non-diseased group OR_X=Observed OR

		OR_T:	1.05	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9
Specificity	Sensitivity	P_N	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X
0.1	0.1	0.1	0.98	0.95	0.91	0.87	0.84	0.80	0.77	0.75	0.72	0.70
0.1	0.1	0.2	0.97	0.94	0.88	0.83	0.79	0.75	0.72	0.69	0.66	0.63
0.1	0.1	0.3	0.96	0.93	0.87	0.82	0.77	0.73	0.70	0.66	0.63	0.61
0.1	0.1	0.4	0.96	0.93	0.87	0.81	0.77	0.72	0.69	0.66	0.63	0.60
0.1	0.1	0.5	0.96	0.93	0.86	0.81	0.76	0.72	0.69	0.66	0.63	0.60
0.1	0.2	0.1	0.98	0.96	0.92	0.88	0.85	0.82	0.79	0.76	0.74	0.72
0.1	0.2	0.2	0.97	0.94	0.89	0.85	0.81	0.77	0.74	0.71	0.68	0.65
0.1	0.2	0.3	0.97	0.94	0.88	0.83	0.79	0.75	0.72	0.69	0.66	0.64
0.1	0.2	0.4	0.97	0.93	0.88	0.83	0.79	0.75	0.72	0.69	0.66	0.63
0.1	0.2	0.5	0.97	0.94	0.88	0.83	0.79	0.76	0.72	0.69	0.67	0.65
0.1	0.3	0.1	0.98	0.96	0.93	0.89	0.86	0.84	0.81	0.78	0.76	0.74
0.1	0.3	0.2	0.97	0.95	0.90	0.86	0.82	0.79	0.76	0.73	0.70	0.68
0.1	0.3	0.3	0.97	0.94	0.89	0.85	0.81	0.77	0.74	0.71	0.69	0.66
0.1	0.3	0.4	0.97	0.94	0.89	0.85	0.81	0.77	0.74	0.72	0.69	0.67
0.1	0.3	0.5	0.97	0.94	0.89	0.85	0.81	0.78	0.75	0.73	0.71	0.69
0.1	0.4	0.1	0.98	0.97	0.93	0.91	0.88	0.85	0.83	0.81	0.79	0.77
0.1	0.4	0.2	0.98	0.95	0.91	0.87	0.84	0.81	0.78	0.75	0.73	0.71
0.1	0.4	0.3	0.97	0.95	0.90	0.86	0.83	0.79	0.77	0.74	0.72	0.69
0.1	0.4	0.4	0.97	0.95	0.90	0.86	0.83	0.80	0.77	0.74	0.72	0.70
0.1	0.4	0.5	0.97	0.95	0.91	0.87	0.84	0.81	0.78	0.76	0.74	0.72
0.1	0.5	0.1	0.99	0.97	0.94	0.92	0.90	0.87	0.85	0.83	0.81	0.80
0.1	0.5	0.2	0.98	0.96	0.92	0.89	0.86	0.83	0.80	0.78	0.76	0.74
0.1	0.5	0.3	0.98	0.95	0.91	0.88	0.85	0.82	0.79	0.77	0.75	0.73
0.1	0.5	0.4	0.98	0.95	0.91	0.88	0.85	0.82	0.80	0.77	0.75	0.73
0.1	0.5	0.5	0.98	0.96	0.92	0.89	0.86	0.83	0.81	0.79	0.77	0.76
0.1	0.6	0.1	0.99	0.98	0.96	0.93	0.91	0.90	0.88	0.86	0.85	0.83
0.1	0.6	0.2	0.98	0.97	0.94	0.91	0.88	0.86	0.84	0.81	0.80	0.78
0.1	0.6	0.3	0.98	0.96	0.93	0.90	0.87	0.84	0.82	0.80	0.78	0.77
0.1	0.6	0.4	0.98	0.96	0.93	0.90	0.87	0.85	0.82	0.81	0.79	0.77
0.1	0.6	0.5	0.98	0.96	0.93	0.90	0.88	0.86	0.84	0.82	0.80	0.79
0.1	0.7	0.1	0.99	0.98	0.97	0.95	0.94	0.92	0.91	0.90	0.89	0.87
0.1	0.7	0.2	0.99	0.97	0.95	0.93	0.91	0.89	0.87	0.86	0.84	0.83
0.1	0.7	0.3	0.98	0.97	0.94	0.92	0.90	0.88	0.86	0.84	0.83	0.81
0.1	0.7	0.4	0.98	0.97	0.94	0.92	0.90	0.88	0.86	0.84	0.83	0.82
0.1	0.7	0.5	0.98	0.97	0.95	0.92	0.90	0.89	0.87	0.86	0.84	0.83
0.1	0.8	0.1	1.00	0.99	0.98	0.97	0.97	0.96	0.95	0.94	0.94	0.93
0.1	0.8	0.2	0.99	0.99	0.97	0.96	0.95	0.94	0.92	0.91	0.90	0.90
0.1	0.8	0.3	0.99	0.98	0.97	0.95	0.94	0.92	0.91	0.90	0.89	0.88
0.1	0.8	0.4	0.99	0.98	0.96	0.95	0.94	0.92	0.91	0.90	0.89	0.88
0.1	0.8	0.5	0.99	0.98	0.97	0.95	0.94	0.93	0.92	0.91	0.90	0.89
0.1	0.9	0.1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.1	0.9	0.2	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.1	0.9	0.3	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.1	0.9	0.4	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.1	0.9	0.5	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.1	0.95	0.1	1.00	1.01	1.01	1.02	1.02	1.03	1.03	1.04	1.04	1.04
0.1	0.95	0.2	1.00	1.01	1.02	1.03	1.04	1.05	1.05	1.06	1.07	1.08
0.1	0.95	0.3	1.01	1.01	1.03	1.04	1.05	1.06	1.07	1.08	1.09	1.10
0.1	0.95	0.4	1.01	1.02	1.03	1.05	1.06	1.07	1.09	1.10	1.11	1.12
0.1	0.95	0.5	1.01	1.02	1.03	1.05	1.06	1.08	1.09	1.10	1.11	1.12
0.1	0.99	0.1	1.00	1.01	1.02	1.03	1.04	1.05	1.06	1.07	1.08	1.09
0.1	0.99	0.2	1.01	1.02	1.04	1.06	1.08	1.09	1.11	1.13	1.15	1.17
0.1	0.99	0.3	1.01	1.03	1.06	1.08	1.11	1.14	1.16	1.19	1.22	1.24
0.1	0.99	0.4	1.02	1.04	1.07	1.11	1.14	1.17	1.21	1.24	1.27	1.31
0.1	0.99	0.5	1.02	1.04	1.09	1.13	1.17	1.21	1.25	1.28	1.32	1.36

Table 3- 6. Effect of non-differential misclassification of categorical data

P_N=the proportion exposed in the non-diseased group

		OR_T:	2.0	2.1	2.2	2.3	2.4	2.5	2.6	2.7	2.8	2.9
Specificity	**Sensitivity**	P_N	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X
0.1	0.1	0.1	0.67	0.65	0.63	0.62	0.60	0.58	0.57	0.55	0.54	0.52
0.1	0.1	0.2	0.61	0.58	0.56	0.54	0.53	0.51	0.49	0.48	0.47	0.45
0.1	0.1	0.3	0.58	0.56	0.54	0.52	0.50	0.49	0.47	0.46	0.45	0.43
0.1	0.1	0.4	0.58	0.55	0.53	0.52	0.50	0.48	0.47	0.45	0.44	0.43
0.1	0.1	0.5	0.58	0.56	0.54	0.52	0.50	0.49	0.48	0.46	0.45	0.44
0.1	0.2	0.1	0.70	0.68	0.66	0.64	0.62	0.61	0.59	0.58	0.56	0.55
0.1	0.2	0.2	0.63	0.61	0.59	0.57	0.56	0.54	0.52	0.51	0.50	0.49
0.1	0.2	0.3	0.61	0.59	0.57	0.55	0.54	0.52	0.51	0.50	0.48	0.47
0.1	0.2	0.4	0.61	0.59	0.57	0.56	0.54	0.53	0.51	0.50	0.49	0.48
0.1	0.2	0.5	0.63	0.61	0.59	0.57	0.56	0.55	0.53	0.52	0.51	0.50
0.1	0.3	0.1	0.72	0.70	0.68	0.67	0.65	0.64	0.62	0.61	0.60	0.58
0.1	0.3	0.2	0.66	0.64	0.62	0.60	0.59	0.57	0.56	0.54	0.53	0.52
0.1	0.3	0.3	0.64	0.62	0.61	0.59	0.57	0.56	0.55	0.53	0.52	0.51
0.1	0.3	0.4	0.65	0.63	0.61	0.60	0.58	0.57	0.56	0.55	0.53	0.52
0.1	0.3	0.5	0.67	0.65	0.63	0.62	0.61	0.59	0.58	0.57	0.56	0.55
0.1	0.4	0.1	0.75	0.73	0.71	0.70	0.68	0.67	0.66	0.64	0.63	0.62
0.1	0.4	0.2	0.69	0.67	0.65	0.63	0.62	0.61	0.59	0.58	0.57	0.56
0.1	0.4	0.3	0.67	0.66	0.64	0.62	0.61	0.60	0.58	0.57	0.56	0.55
0.1	0.4	0.4	0.68	0.67	0.65	0.64	0.62	0.61	0.60	0.59	0.58	0.57
0.1	0.4	0.5	0.70	0.69	0.67	0.66	0.65	0.64	0.63	0.62	0.61	0.60
0.1	0.5	0.1	0.78	0.76	0.75	0.73	0.72	0.71	0.70	0.68	0.67	0.66
0.1	0.5	0.2	0.72	0.70	0.69	0.67	0.66	0.65	0.63	0.62	0.61	0.60
0.1	0.5	0.3	0.71	0.69	0.68	0.66	0.65	0.64	0.63	0.61	0.60	0.59
0.1	0.5	0.4	0.72	0.70	0.69	0.68	0.66	0.65	0.64	0.63	0.62	0.61
0.1	0.5	0.5	0.74	0.73	0.71	0.70	0.69	0.68	0.67	0.67	0.66	0.65
0.1	0.6	0.1	0.82	0.80	0.79	0.78	0.77	0.75	0.74	0.73	0.72	0.71
0.1	0.6	0.2	0.76	0.75	0.73	0.72	0.71	0.69	0.68	0.67	0.66	0.65
0.1	0.6	0.3	0.75	0.73	0.72	0.71	0.70	0.68	0.67	0.66	0.66	0.65
0.1	0.6	0.4	0.76	0.74	0.73	0.72	0.71	0.70	0.69	0.68	0.67	0.67
0.1	0.6	0.5	0.78	0.77	0.76	0.75	0.74	0.73	0.72	0.71	0.70	0.70
0.1	0.7	0.1	0.86	0.85	0.84	0.83	0.82	0.81	0.81	0.80	0.79	0.78
0.1	0.7	0.2	0.81	0.80	0.79	0.78	0.77	0.76	0.75	0.74	0.73	0.72
0.1	0.7	0.3	0.80	0.79	0.78	0.77	0.76	0.75	0.74	0.73	0.72	0.71
0.1	0.7	0.4	0.80	0.79	0.78	0.77	0.76	0.76	0.75	0.74	0.73	0.73
0.1	0.7	0.5	0.82	0.81	0.80	0.79	0.79	0.78	0.77	0.77	0.76	0.76
0.1	0.8	0.1	0.92	0.92	0.91	0.90	0.90	0.89	0.89	0.88	0.88	0.87
0.1	0.8	0.2	0.89	0.88	0.87	0.86	0.86	0.85	0.84	0.84	0.83	0.82
0.1	0.8	0.3	0.87	0.86	0.86	0.85	0.84	0.84	0.83	0.82	0.82	0.81
0.1	0.8	0.4	0.87	0.87	0.86	0.85	0.84	0.84	0.83	0.83	0.82	0.82
0.1	0.8	0.5	0.88	0.88	0.87	0.86	0.86	0.85	0.85	0.84	0.84	0.84
0.1	0.9	0.1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.1	0.9	0.2	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.1	0.9	0.3	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.1	0.9	0.4	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.1	0.9	0.5	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.1	0.95	0.1	1.05	1.05	1.06	1.06	1.07	1.07	1.08	1.08	1.09	1.09
0.1	0.95	0.2	1.09	1.10	1.10	1.11	1.12	1.13	1.13	1.14	1.15	1.15
0.1	0.95	0.3	1.11	1.12	1.13	1.14	1.15	1.16	1.17	1.18	1.18	1.19
0.1	0.95	0.4	1.13	1.14	1.15	1.16	1.17	1.18	1.19	1.19	1.20	1.21
0.1	0.95	0.5	1.14	1.15	1.15	1.16	1.17	1.18	1.19	1.20	1.20	1.21
0.1	0.99	0.1	1.10	1.11	1.12	1.13	1.14	1.14	1.15	1.16	1.17	1.18
0.1	0.99	0.2	1.19	1.20	1.22	1.24	1.26	1.28	1.29	1.31	1.33	1.35
0.1	0.99	0.3	1.27	1.29	1.32	1.34	1.37	1.39	1.42	1.44	1.47	1.49
0.1	0.99	0.4	1.34	1.37	1.40	1.43	1.46	1.49	1.52	1.55	1.58	1.61
0.1	0.99	0.5	1.40	1.43	1.47	1.50	1.54	1.57	1.60	1.64	1.67	1.70

Table 3- 6. Effect of non-differential misclassification of categorical data
P_N=the proportion exposed in the non-diseased group

		OR_T:	3.0	3.1	3.2	3.3	3.4	3.5
Specificity	**Sensitivity**	**P_N**	**OR_X**	**OR_X**	**OR_X**	**OR_X**	**OR_X**	**OR_X**
0.1	0.1	0.1	0.51	0.50	0.49	0.48	0.47	0.46
0.1	0.1	0.2	0.44	0.43	0.42	0.41	0.40	0.39
0.1	0.1	0.3	0.42	0.41	0.40	0.39	0.38	0.37
0.1	0.1	0.4	0.42	0.41	0.40	0.39	0.38	0.37
0.1	0.1	0.5	0.43	0.42	0.41	0.40	0.39	0.38
0.1	0.2	0.1	0.54	0.53	0.52	0.51	0.50	0.49
0.1	0.2	0.2	0.47	0.46	0.45	0.44	0.43	0.42
0.1	0.2	0.3	0.46	0.45	0.44	0.43	0.42	0.41
0.1	0.2	0.4	0.47	0.46	0.45	0.44	0.43	0.43
0.1	0.2	0.5	0.49	0.48	0.47	0.47	0.46	0.45
0.1	0.3	0.1	0.57	0.56	0.55	0.54	0.53	0.52
0.1	0.3	0.2	0.51	0.50	0.49	0.48	0.47	0.46
0.1	0.3	0.3	0.50	0.49	0.48	0.47	0.46	0.46
0.1	0.3	0.4	0.52	0.51	0.50	0.49	0.48	0.48
0.1	0.3	0.5	0.55	0.54	0.53	0.52	0.52	0.51
0.1	0.4	0.1	0.61	0.60	0.59	0.58	0.57	0.56
0.1	0.4	0.2	0.55	0.54	0.53	0.52	0.51	0.50
0.1	0.4	0.3	0.54	0.53	0.52	0.52	0.51	0.50
0.1	0.4	0.4	0.56	0.55	0.54	0.54	0.53	0.52
0.1	0.4	0.5	0.60	0.59	0.58	0.57	0.57	0.56
0.1	0.5	0.1	0.65	0.64	0.63	0.62	0.61	0.61
0.1	0.5	0.2	0.59	0.58	0.57	0.56	0.55	0.55
0.1	0.5	0.3	0.59	0.58	0.57	0.56	0.55	0.55
0.1	0.5	0.4	0.61	0.60	0.59	0.59	0.58	0.57
0.1	0.5	0.5	0.64	0.64	0.63	0.62	0.62	0.61
0.1	0.6	0.1	0.70	0.70	0.69	0.68	0.67	0.66
0.1	0.6	0.2	0.64	0.63	0.63	0.62	0.61	0.60
0.1	0.6	0.3	0.64	0.63	0.62	0.62	0.61	0.60
0.1	0.6	0.4	0.66	0.65	0.64	0.64	0.63	0.63
0.1	0.6	0.5	0.69	0.69	0.68	0.68	0.67	0.67
0.1	0.7	0.1	0.77	0.77	0.76	0.75	0.74	0.74
0.1	0.7	0.2	0.71	0.71	0.70	0.69	0.69	0.68
0.1	0.7	0.3	0.71	0.70	0.69	0.69	0.68	0.68
0.1	0.7	0.4	0.72	0.72	0.71	0.70	0.70	0.70
0.1	0.7	0.5	0.75	0.75	0.74	0.74	0.73	0.73
0.1	0.8	0.1	0.87	0.86	0.86	0.85	0.85	0.84
0.1	0.8	0.2	0.82	0.81	0.81	0.80	0.80	0.79
0.1	0.8	0.3	0.81	0.80	0.80	0.79	0.79	0.78
0.1	0.8	0.4	0.81	0.81	0.81	0.80	0.80	0.79
0.1	0.8	0.5	0.83	0.83	0.83	0.82	0.82	0.82
0.1	0.9	0.1	1.00	1.00	1.00	1.00	1.00	1.00
0.1	0.9	0.2	1.00	1.00	1.00	1.00	1.00	1.00
0.1	0.9	0.3	1.00	1.00	1.00	1.00	1.00	1.00
0.1	0.9	0.4	1.00	1.00	1.00	1.00	1.00	1.00
0.1	0.9	0.5	1.00	1.00	1.00	1.00	1.00	1.00
0.1	0.95	0.1	1.09	1.10	1.10	1.11	1.11	1.12
0.1	0.95	0.2	1.16	1.17	1.17	1.18	1.19	1.19
0.1	0.95	0.3	1.20	1.21	1.21	1.22	1.23	1.23
0.1	0.95	0.4	1.22	1.22	1.23	1.24	1.24	1.25
0.1	0.95	0.5	1.22	1.22	1.23	1.23	1.24	1.25
0.1	0.99	0.1	1.19	1.20	1.21	1.22	1.23	1.24
0.1	0.99	0.2	1.36	1.38	1.40	1.42	1.43	1.45
0.1	0.99	0.3	1.52	1.54	1.56	1.59	1.61	1.63
0.1	0.99	0.4	1.64	1.67	1.70	1.73	1.75	1.78
0.1	0.99	0.5	1.73	1.76	1.79	1.82	1.85	1.88

Table 3-6. Effect of non-differential misclassification of categorical data continued

P_N=the proportion exposed in the non-diseased group OR_X=Observed OR

		OR_T	1.05	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9
Specificity	**Sensitivity**	P_N	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X
0.2	0.1	0.1	0.98	0.97	0.94	0.91	0.89	0.86	0.84	0.82	0.80	0.78
0.2	0.1	0.2	0.98	0.95	0.91	0.87	0.84	0.80	0.77	0.75	0.72	0.70
0.2	0.1	0.3	0.97	0.94	0.89	0.85	0.81	0.77	0.74	0.71	0.68	0.66
0.2	0.1	0.4	0.97	0.94	0.88	0.84	0.79	0.76	0.72	0.69	0.66	0.64
0.2	0.1	0.5	0.97	0.93	0.88	0.83	0.79	0.75	0.72	0.69	0.66	0.63
0.2	0.2	0.1	0.99	0.97	0.95	0.92	0.90	0.88	0.86	0.84	0.82	0.80
0.2	0.2	0.2	0.98	0.96	0.92	0.88	0.85	0.82	0.80	0.77	0.75	0.73
0.2	0.2	0.3	0.97	0.95	0.91	0.87	0.83	0.80	0.77	0.74	0.72	0.69
0.2	0.2	0.4	0.97	0.95	0.90	0.86	0.82	0.79	0.76	0.73	0.70	0.68
0.2	0.2	0.5	0.97	0.94	0.90	0.85	0.82	0.79	0.76	0.73	0.71	0.69
0.2	0.3	0.1	0.99	0.98	0.95	0.93	0.91	0.89	0.88	0.86	0.84	0.83
0.2	0.3	0.2	0.98	0.96	0.93	0.90	0.87	0.85	0.82	0.80	0.78	0.76
0.2	0.3	0.3	0.98	0.96	0.92	0.88	0.85	0.82	0.80	0.77	0.75	0.73
0.2	0.3	0.4	0.98	0.95	0.91	0.88	0.84	0.81	0.79	0.77	0.74	0.72
0.2	0.3	0.5	0.98	0.95	0.91	0.88	0.85	0.82	0.79	0.77	0.75	0.73
0.2	0.4	0.1	0.99	0.98	0.96	0.94	0.93	0.91	0.90	0.88	0.87	0.85
0.2	0.4	0.2	0.98	0.97	0.94	0.92	0.89	0.87	0.85	0.83	0.81	0.79
0.2	0.4	0.3	0.98	0.96	0.93	0.90	0.87	0.85	0.83	0.81	0.79	0.77
0.2	0.4	0.4	0.98	0.96	0.93	0.90	0.87	0.84	0.82	0.80	0.78	0.77
0.2	0.4	0.5	0.98	0.96	0.93	0.90	0.87	0.85	0.83	0.81	0.79	0.78
0.2	0.5	0.1	0.99	0.99	0.97	0.96	0.94	0.93	0.92	0.91	0.90	0.89
0.2	0.5	0.2	0.99	0.98	0.95	0.93	0.91	0.90	0.88	0.86	0.85	0.83
0.2	0.5	0.3	0.99	0.97	0.94	0.92	0.90	0.88	0.86	0.84	0.83	0.81
0.2	0.5	0.4	0.98	0.97	0.94	0.92	0.89	0.87	0.86	0.84	0.82	0.81
0.2	0.5	0.5	0.98	0.97	0.94	0.92	0.90	0.88	0.86	0.85	0.83	0.82
0.2	0.6	0.1	0.99	0.99	0.98	0.97	0.96	0.95	0.94	0.93	0.93	0.92
0.2	0.6	0.2	0.99	0.98	0.97	0.95	0.94	0.92	0.91	0.90	0.89	0.88
0.2	0.6	0.3	0.99	0.98	0.96	0.94	0.93	0.91	0.90	0.88	0.87	0.86
0.2	0.6	0.4	0.99	0.98	0.96	0.94	0.92	0.91	0.89	0.88	0.87	0.86
0.2	0.6	0.5	0.99	0.98	0.96	0.94	0.92	0.91	0.90	0.89	0.88	0.87
0.2	0.7	0.1	1.00	0.99	0.99	0.98	0.98	0.97	0.97	0.97	0.96	0.96
0.2	0.7	0.2	1.00	0.99	0.98	0.97	0.97	0.96	0.95	0.95	0.94	0.93
0.2	0.7	0.3	0.99	0.99	0.98	0.97	0.96	0.95	0.94	0.93	0.93	0.92
0.2	0.7	0.4	0.99	0.99	0.98	0.97	0.96	0.95	0.94	0.93	0.92	0.92
0.2	0.7	0.5	0.99	0.99	0.98	0.97	0.96	0.95	0.94	0.93	0.93	0.92
0.2	0.8	0.1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.2	0.8	0.2	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.2	0.8	0.3	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.2	0.8	0.4	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.2	0.8	0.5	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.2	0.9	0.1	1.00	1.01	1.01	1.02	1.02	1.03	1.03	1.04	1.04	1.05
0.2	0.9	0.2	1.01	1.01	1.02	1.03	1.04	1.05	1.06	1.07	1.08	1.09
0.2	0.9	0.3	1.01	1.01	1.03	1.04	1.06	1.07	1.08	1.09	1.10	1.12
0.2	0.9	0.4	1.01	1.02	1.03	1.05	1.06	1.08	1.09	1.11	1.12	1.13
0.2	0.9	0.5	1.01	1.02	1.04	1.05	1.07	1.08	1.10	1.11	1.12	1.14
0.2	0.95	0.1	1.00	1.01	1.02	1.03	1.04	1.04	1.05	1.06	1.07	1.08
0.2	0.95	0.2	1.01	1.02	1.03	1.05	1.07	1.08	1.10	1.11	1.13	1.15
0.2	0.95	0.3	1.01	1.02	1.05	1.07	1.09	1.11	1.14	1.16	1.18	1.20
0.2	0.95	0.4	1.01	1.03	1.06	1.09	1.11	1.14	1.17	1.19	1.21	1.24
0.2	0.95	0.5	1.02	1.03	1.07	1.10	1.13	1.16	1.18	1.21	1.24	1.26
0.2	0.99	0.1	1.01	1.01	1.02	1.03	1.05	1.06	1.07	1.08	1.09	1.10
0.2	0.99	0.2	1.01	1.02	1.04	1.07	1.09	1.11	1.13	1.16	1.18	1.20
0.2	0.99	0.3	1.02	1.03	1.06	1.10	1.13	1.16	1.19	1.22	1.26	1.29
0.2	0.99	0.4	1.02	1.04	1.08	1.12	1.17	1.21	1.25	1.29	1.33	1.37
0.2	0.99	0.5	1.03	1.05	1.10	1.15	1.20	1.25	1.29	1.34	1.39	1.44

Table 3-6. Effect of non-differential misclassification of categorical data continued
P_N=the proportion exposed in the non-diseased group

		OR_T	2.0	2.1	2.2	2.3	2.4	2.5	2.6	2.7	2.8	2.9
Specificity	**Sensitivity**	P_N	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X
0.2	0.1	0.1	0.76	0.74	0.73	0.71	0.69	0.68	0.67	0.65	0.64	0.63
0.2	0.1	0.2	0.67	0.65	0.63	0.62	0.60	0.58	0.57	0.55	0.54	0.53
0.2	0.1	0.3	0.63	0.61	0.59	0.57	0.56	0.54	0.53	0.51	0.50	0.49
0.2	0.1	0.4	0.62	0.59	0.57	0.56	0.54	0.52	0.51	0.50	0.48	0.47
0.2	0.1	0.5	0.61	0.59	0.57	0.55	0.54	0.52	0.51	0.50	0.49	0.47
0.2	0.2	0.1	0.79	0.77	0.75	0.74	0.73	0.71	0.70	0.69	0.67	0.66
0.2	0.2	0.2	0.71	0.69	0.67	0.65	0.64	0.62	0.61	0.59	0.58	0.57
0.2	0.2	0.3	0.67	0.65	0.63	0.62	0.60	0.59	0.57	0.56	0.55	0.54
0.2	0.2	0.4	0.66	0.64	0.63	0.61	0.59	0.58	0.57	0.56	0.54	0.53
0.2	0.2	0.5	0.67	0.65	0.63	0.62	0.60	0.59	0.58	0.57	0.56	0.55
0.2	0.3	0.1	0.81	0.80	0.78	0.77	0.76	0.75	0.73	0.72	0.71	0.70
0.2	0.3	0.2	0.74	0.72	0.71	0.69	0.68	0.66	0.65	0.64	0.63	0.62
0.2	0.3	0.3	0.71	0.69	0.68	0.66	0.65	0.64	0.62	0.61	0.60	0.59
0.2	0.3	0.4	0.71	0.69	0.67	0.66	0.65	0.63	0.62	0.61	0.60	0.59
0.2	0.3	0.5	0.72	0.70	0.69	0.67	0.66	0.65	0.64	0.63	0.62	0.61
0.2	0.4	0.1	0.84	0.83	0.82	0.81	0.80	0.78	0.77	0.76	0.75	0.75
0.2	0.4	0.2	0.78	0.76	0.75	0.73	0.72	0.71	0.70	0.69	0.68	0.67
0.2	0.4	0.3	0.75	0.74	0.72	0.71	0.70	0.69	0.67	0.66	0.65	0.65
0.2	0.4	0.4	0.75	0.74	0.72	0.71	0.70	0.69	0.68	0.67	0.66	0.65
0.2	0.4	0.5	0.76	0.75	0.74	0.73	0.72	0.71	0.70	0.69	0.68	0.67
0.2	0.5	0.1	0.87	0.86	0.85	0.85	0.84	0.83	0.82	0.81	0.80	0.80
0.2	0.5	0.2	0.82	0.81	0.80	0.78	0.77	0.76	0.75	0.74	0.73	0.73
0.2	0.5	0.3	0.80	0.79	0.77	0.76	0.75	0.74	0.73	0.72	0.71	0.71
0.2	0.5	0.4	0.80	0.78	0.77	0.76	0.75	0.74	0.74	0.73	0.72	0.71
0.2	0.5	0.5	0.81	0.80	0.79	0.78	0.77	0.76	0.75	0.75	0.74	0.73
0.2	0.6	0.1	0.91	0.90	0.90	0.89	0.88	0.88	0.87	0.86	0.86	0.85
0.2	0.6	0.2	0.87	0.86	0.85	0.84	0.83	0.82	0.82	0.81	0.80	0.80
0.2	0.6	0.3	0.85	0.84	0.83	0.82	0.81	0.81	0.80	0.79	0.79	0.78
0.2	0.6	0.4	0.85	0.84	0.83	0.82	0.81	0.81	0.80	0.79	0.79	0.78
0.2	0.6	0.5	0.86	0.85	0.84	0.83	0.83	0.82	0.82	0.81	0.81	0.80
0.2	0.7	0.1	0.95	0.95	0.94	0.94	0.94	0.93	0.93	0.93	0.92	0.92
0.2	0.7	0.2	0.93	0.92	0.92	0.91	0.91	0.90	0.90	0.89	0.89	0.88
0.2	0.7	0.3	0.91	0.91	0.90	0.90	0.89	0.89	0.88	0.88	0.87	0.87
0.2	0.7	0.4	0.91	0.91	0.90	0.90	0.89	0.89	0.88	0.88	0.88	0.87
0.2	0.7	0.5	0.92	0.91	0.91	0.90	0.90	0.89	0.89	0.89	0.88	0.88
0.2	0.8	0.1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.2	0.8	0.2	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.2	0.8	0.3	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.2	0.8	0.4	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.2	0.8	0.5	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.2	0.9	0.1	1.06	1.06	1.07	1.07	1.08	1.08	1.09	1.09	1.10	1.10
0.2	0.9	0.2	1.10	1.11	1.11	1.12	1.13	1.14	1.15	1.16	1.16	1.17
0.2	0.9	0.3	1.13	1.14	1.15	1.16	1.17	1.18	1.19	1.19	1.20	1.21
0.2	0.9	0.4	1.14	1.15	1.16	1.18	1.19	1.19	1.20	1.21	1.22	1.23
0.2	0.9	0.5	1.15	1.16	1.17	1.18	1.19	1.20	1.20	1.21	1.22	1.23
0.2	0.95	0.1	1.09	1.10	1.10	1.11	1.12	1.13	1.14	1.15	1.15	1.16
0.2	0.95	0.2	1.16	1.18	1.19	1.21	1.22	1.23	1.25	1.26	1.28	1.29
0.2	0.95	0.3	1.22	1.24	1.26	1.28	1.30	1.32	1.33	1.35	1.37	1.39
0.2	0.95	0.4	1.26	1.28	1.31	1.33	1.35	1.37	1.39	1.41	1.43	1.45
0.2	0.95	0.5	1.29	1.31	1.33	1.35	1.38	1.40	1.42	1.43	1.45	1.47
0.2	0.99	0.1	1.11	1.13	1.14	1.15	1.16	1.17	1.18	1.19	1.21	1.22
0.2	0.99	0.2	1.22	1.24	1.26	1.29	1.31	1.33	1.35	1.37	1.39	1.42
0.2	0.99	0.3	1.32	1.35	1.38	1.41	1.44	1.47	1.50	1.53	1.56	1.59
0.2	0.99	0.4	1.41	1.45	1.48	1.52	1.56	1.60	1.64	1.68	1.71	1.75
0.2	0.99	0.5	1.48	1.53	1.57	1.62	1.66	1.71	1.75	1.79	1.84	1.88

Table 3-6. Effect of non-differential misclassification of categorical data continued
P_N=the proportion exposed in the non-diseased group

		OR_T	3.0	3.1	3.2	3.3	3.4	3.5	3.6	3.7	3.8	3.9
Specificity	Sensitivity	P_N	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X
0.2	0.1	0.1	0.62	0.61	0.59	0.58	0.57	0.56	0.55	0.55	0.54	0.53
0.2	0.1	0.2	0.52	0.50	0.49	0.48	0.47	0.46	0.45	0.45	0.44	0.43
0.2	0.1	0.3	0.48	0.46	0.45	0.44	0.43	0.43	0.42	0.41	0.40	0.39
0.2	0.1	0.4	0.46	0.45	0.44	0.43	0.42	0.41	0.41	0.40	0.39	0.39
0.2	0.1	0.5	0.46	0.45	0.44	0.44	0.43	0.42	0.41	0.41	0.40	0.39
0.2	0.2	0.1	0.65	0.64	0.63	0.62	0.61	0.60	0.59	0.59	0.58	0.57
0.2	0.2	0.2	0.56	0.55	0.54	0.53	0.52	0.51	0.50	0.49	0.49	0.48
0.2	0.2	0.3	0.53	0.52	0.51	0.50	0.49	0.48	0.47	0.47	0.46	0.45
0.2	0.2	0.4	0.52	0.51	0.51	0.50	0.49	0.48	0.47	0.47	0.46	0.45
0.2	0.2	0.5	0.54	0.53	0.52	0.51	0.51	0.50	0.49	0.49	0.48	0.48
0.2	0.3	0.1	0.69	0.68	0.67	0.66	0.66	0.65	0.64	0.63	0.62	0.62
0.2	0.3	0.2	0.61	0.60	0.59	0.58	0.57	0.56	0.55	0.54	0.54	0.53
0.2	0.3	0.3	0.58	0.57	0.56	0.55	0.55	0.54	0.53	0.52	0.52	0.51
0.2	0.3	0.4	0.58	0.57	0.57	0.56	0.55	0.55	0.54	0.53	0.53	0.52
0.2	0.3	0.5	0.60	0.60	0.59	0.58	0.58	0.57	0.57	0.56	0.55	0.55
0.2	0.4	0.1	0.74	0.73	0.72	0.71	0.70	0.70	0.69	0.68	0.67	0.67
0.2	0.4	0.2	0.66	0.65	0.64	0.63	0.62	0.62	0.61	0.60	0.60	0.59
0.2	0.4	0.3	0.64	0.63	0.62	0.61	0.61	0.60	0.59	0.59	0.58	0.57
0.2	0.4	0.4	0.64	0.64	0.63	0.62	0.62	0.61	0.60	0.60	0.59	0.59
0.2	0.4	0.5	0.67	0.66	0.65	0.65	0.64	0.64	0.63	0.63	0.62	0.62
0.2	0.5	0.1	0.79	0.78	0.77	0.77	0.76	0.75	0.75	0.74	0.73	0.73
0.2	0.5	0.2	0.72	0.71	0.70	0.70	0.69	0.68	0.68	0.67	0.66	0.66
0.2	0.5	0.3	0.70	0.69	0.69	0.68	0.67	0.67	0.66	0.66	0.65	0.65
0.2	0.5	0.4	0.71	0.70	0.69	0.69	0.68	0.68	0.67	0.67	0.66	0.66
0.2	0.5	0.5	0.73	0.72	0.72	0.71	0.71	0.70	0.70	0.70	0.69	0.69
0.2	0.6	0.1	0.85	0.84	0.84	0.83	0.82	0.82	0.81	0.81	0.81	0.80
0.2	0.6	0.2	0.79	0.78	0.78	0.77	0.77	0.76	0.76	0.75	0.75	0.74
0.2	0.6	0.3	0.77	0.77	0.76	0.76	0.75	0.75	0.74	0.74	0.73	0.73
0.2	0.6	0.4	0.78	0.77	0.77	0.76	0.76	0.75	0.75	0.75	0.74	0.74
0.2	0.6	0.5	0.80	0.79	0.79	0.78	0.78	0.78	0.77	0.77	0.77	0.76
0.2	0.7	0.1	0.92	0.91	0.91	0.91	0.90	0.90	0.90	0.89	0.89	0.89
0.2	0.7	0.2	0.88	0.88	0.87	0.87	0.86	0.86	0.86	0.86	0.85	0.85
0.2	0.7	0.3	0.87	0.86	0.86	0.86	0.85	0.85	0.85	0.84	0.84	0.84
0.2	0.7	0.4	0.87	0.87	0.86	0.86	0.86	0.85	0.85	0.85	0.85	0.84
0.2	0.7	0.5	0.88	0.88	0.87	0.87	0.87	0.87	0.86	0.86	0.86	0.86
0.2	0.8	0.1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.2	0.8	0.2	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.2	0.8	0.3	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.2	0.8	0.4	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.2	0.8	0.5	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.2	0.9	0.1	1.11	1.11	1.12	1.12	1.12	1.13	1.13	1.14	1.14	1.15
0.2	0.9	0.2	1.18	1.18	1.19	1.20	1.21	1.21	1.22	1.23	1.23	1.24
0.2	0.9	0.3	1.22	1.23	1.24	1.24	1.25	1.26	1.27	1.27	1.28	1.29
0.2	0.9	0.4	1.24	1.25	1.25	1.26	1.27	1.27	1.28	1.29	1.29	1.30
0.2	0.9	0.5	1.24	1.24	1.25	1.26	1.26	1.27	1.27	1.28	1.28	1.29
0.2	0.95	0.1	1.17	1.18	1.19	1.19	1.20	1.21	1.22	1.23	1.23	1.24
0.2	0.95	0.2	1.30	1.32	1.33	1.34	1.36	1.37	1.38	1.40	1.41	1.42
0.2	0.95	0.3	1.40	1.42	1.44	1.45	1.47	1.48	1.50	1.51	1.53	1.54
0.2	0.95	0.4	1.47	1.48	1.50	1.52	1.53	1.55	1.57	1.58	1.60	1.61
0.2	0.95	0.5	1.49	1.51	1.52	1.54	1.56	1.57	1.59	1.60	1.62	1.63
0.2	0.99	0.1	1.23	1.24	1.25	1.26	1.27	1.28	1.30	1.31	1.32	1.33
0.2	0.99	0.2	1.44	1.46	1.48	1.50	1.52	1.54	1.56	1.59	1.61	1.63
0.2	0.99	0.3	1.62	1.65	1.68	1.71	1.74	1.77	1.80	1.83	1.86	1.89
0.2	0.99	0.4	1.79	1.83	1.86	1.90	1.94	1.97	2.01	2.04	2.08	2.11
0.2	0.99	0.5	1.92	1.96	2.01	2.05	2.09	2.13	2.17	2.21	2.25	2.29

Table 3-6. Effect of non-differential misclassification of categorical data continued

P_N=the proportion exposed in the non-diseased grou OR_X=Observed OR

		OR_T	1.05	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9
Specificity	Sensitivity	P_N	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X
0.3	0.1	0.1	0.99	0.98	0.96	0.93	0.91	0.90	0.88	0.86	0.84	0.83
0.3	0.1	0.2	0.98	0.96	0.93	0.90	0.87	0.84	0.81	0.79	0.77	0.74
0.3	0.1	0.3	0.98	0.95	0.91	0.87	0.84	0.80	0.77	0.75	0.72	0.70
0.3	0.1	0.4	0.97	0.95	0.90	0.86	0.82	0.78	0.75	0.72	0.70	0.67
0.3	0.1	0.5	0.97	0.94	0.89	0.85	0.81	0.77	0.74	0.71	0.69	0.66
0.3	0.2	0.1	0.99	0.98	0.96	0.94	0.93	0.91	0.90	0.88	0.87	0.85
0.3	0.2	0.2	0.98	0.97	0.94	0.91	0.89	0.86	0.84	0.82	0.80	0.78
0.3	0.2	0.3	0.98	0.96	0.92	0.89	0.86	0.83	0.81	0.78	0.76	0.74
0.3	0.2	0.4	0.98	0.95	0.91	0.88	0.85	0.82	0.79	0.77	0.75	0.73
0.3	0.2	0.5	0.98	0.95	0.91	0.88	0.84	0.81	0.79	0.77	0.74	0.73
0.3	0.3	0.1	0.99	0.98	0.97	0.95	0.94	0.93	0.91	0.90	0.89	0.88
0.3	0.3	0.2	0.99	0.97	0.95	0.93	0.91	0.89	0.87	0.85	0.83	0.82
0.3	0.3	0.3	0.98	0.97	0.94	0.91	0.89	0.86	0.84	0.82	0.80	0.79
0.3	0.3	0.4	0.98	0.96	0.93	0.90	0.88	0.85	0.83	0.81	0.79	0.78
0.3	0.3	0.5	0.98	0.96	0.93	0.90	0.88	0.85	0.83	0.81	0.79	0.78
0.3	0.4	0.1	0.99	0.99	0.98	0.97	0.95	0.94	0.93	0.92	0.91	0.91
0.3	0.4	0.2	0.99	0.98	0.96	0.94	0.93	0.91	0.90	0.88	0.87	0.86
0.3	0.4	0.3	0.99	0.97	0.95	0.93	0.91	0.89	0.88	0.86	0.85	0.83
0.3	0.4	0.4	0.99	0.97	0.95	0.92	0.90	0.89	0.87	0.85	0.84	0.82
0.3	0.4	0.5	0.99	0.97	0.95	0.92	0.90	0.89	0.87	0.86	0.84	0.83
0.3	0.5	0.1	1.00	0.99	0.98	0.98	0.97	0.96	0.95	0.95	0.94	0.93
0.3	0.5	0.2	0.99	0.99	0.97	0.96	0.95	0.94	0.93	0.92	0.91	0.90
0.3	0.5	0.3	0.99	0.98	0.97	0.95	0.94	0.92	0.91	0.90	0.89	0.88
0.3	0.5	0.4	0.99	0.98	0.96	0.95	0.93	0.92	0.91	0.90	0.89	0.88
0.3	0.5	0.5	0.99	0.98	0.96	0.95	0.93	0.92	0.91	0.90	0.89	0.88
0.3	0.6	0.1	1.00	1.00	0.99	0.99	0.98	0.98	0.98	0.97	0.97	0.97
0.3	0.6	0.2	1.00	0.99	0.99	0.98	0.97	0.97	0.96	0.96	0.95	0.95
0.3	0.6	0.3	1.00	0.99	0.98	0.97	0.97	0.96	0.95	0.95	0.94	0.94
0.3	0.6	0.4	0.99	0.99	0.98	0.97	0.96	0.96	0.95	0.94	0.94	0.93
0.3	0.6	0.5	0.99	0.99	0.98	0.97	0.96	0.96	0.95	0.95	0.94	0.93
0.3	0.7	0.1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.3	0.7	0.2	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.3	0.7	0.3	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.3	0.7	0.4	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.3	0.7	0.5	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.3	0.8	0.1	1.00	1.00	1.01	1.01	1.02	1.02	1.03	1.03	1.03	1.04
0.3	0.8	0.2	1.00	1.01	1.02	1.02	1.03	1.04	1.04	1.05	1.06	1.06
0.3	0.8	0.3	1.01	1.01	1.02	1.03	1.04	1.05	1.06	1.06	1.07	1.08
0.3	0.8	0.4	1.01	1.01	1.02	1.03	1.04	1.05	1.06	1.07	1.08	1.09
0.3	0.8	0.5	1.01	1.01	1.02	1.04	1.05	1.06	1.06	1.07	1.08	1.09
0.3	0.9	0.1	1.00	1.01	1.02	1.03	1.04	1.04	1.05	1.06	1.07	1.08
0.3	0.9	0.2	1.01	1.02	1.03	1.05	1.06	1.08	1.10	1.11	1.13	1.14
0.3	0.9	0.3	1.01	1.02	1.04	1.07	1.09	1.11	1.13	1.15	1.17	1.19
0.3	0.9	0.4	1.01	1.03	1.05	1.08	1.10	1.13	1.15	1.17	1.20	1.22
0.3	0.9	0.5	1.02	1.03	1.06	1.09	1.11	1.14	1.16	1.19	1.21	1.23
0.3	0.95	0.1	1.01	1.01	1.02	1.03	1.04	1.06	1.07	1.08	1.09	1.10
0.3	0.95	0.2	1.01	1.02	1.04	1.06	1.08	1.10	1.13	1.15	1.17	1.19
0.3	0.95	0.3	1.02	1.03	1.06	1.09	1.12	1.15	1.17	1.20	1.23	1.26
0.3	0.95	0.4	1.02	1.04	1.07	1.11	1.14	1.18	1.21	1.25	1.28	1.31
0.3	0.95	0.5	1.02	1.04	1.08	1.12	1.16	1.20	1.24	1.28	1.31	1.35
0.3	0.99	0.1	1.01	1.01	1.03	1.04	1.05	1.07	1.08	1.09	1.11	1.12
0.3	0.99	0.2	1.01	1.03	1.05	1.08	1.10	1.13	1.15	1.18	1.20	1.23
0.3	0.99	0.3	1.02	1.04	1.07	1.11	1.14	1.18	1.22	1.25	1.29	1.32
0.3	0.99	0.4	1.02	1.05	1.09	1.14	1.18	1.23	1.27	1.32	1.36	1.41
0.3	0.99	0.5	1.03	1.06	1.11	1.16	1.22	1.27	1.33	1.38	1.43	1.48

Table 3-6. Effect of non-differential misclassification of categorical data continued

P_N=the proportion exposed in the non-diseased group

		OR_T	2.0	2.1	2.2	2.3	2.4	2.5	2.6	2.7	2.8	2.9
Specificity	Sensitivity	P_N	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X
0.3	0.1	0.1	0.81	0.80	0.78	0.77	0.76	0.74	0.73	0.72	0.71	0.70
0.3	0.1	0.2	0.72	0.71	0.69	0.67	0.66	0.64	0.63	0.61	0.60	0.59
0.3	0.1	0.3	0.68	0.66	0.64	0.62	0.60	0.59	0.57	0.56	0.55	0.54
0.3	0.1	0.4	0.65	0.63	0.61	0.60	0.58	0.57	0.55	0.54	0.53	0.51
0.3	0.1	0.5	0.64	0.62	0.61	0.59	0.57	0.56	0.55	0.53	0.52	0.51
0.3	0.2	0.1	0.84	0.83	0.81	0.80	0.79	0.78	0.77	0.76	0.75	0.74
0.3	0.2	0.2	0.76	0.75	0.73	0.71	0.70	0.69	0.67	0.66	0.65	0.64
0.3	0.2	0.3	0.72	0.71	0.69	0.67	0.66	0.65	0.63	0.62	0.61	0.60
0.3	0.2	0.4	0.71	0.69	0.67	0.66	0.65	0.63	0.62	0.61	0.60	0.59
0.3	0.2	0.5	0.71	0.69	0.68	0.66	0.65	0.64	0.63	0.62	0.61	0.60
0.3	0.3	0.1	0.87	0.86	0.85	0.84	0.83	0.82	0.81	0.80	0.79	0.78
0.3	0.3	0.2	0.80	0.79	0.77	0.76	0.75	0.74	0.73	0.72	0.71	0.70
0.3	0.3	0.3	0.77	0.76	0.74	0.73	0.72	0.70	0.69	0.68	0.67	0.66
0.3	0.3	0.4	0.76	0.75	0.73	0.72	0.71	0.70	0.69	0.68	0.67	0.66
0.3	0.3	0.5	0.76	0.75	0.74	0.73	0.72	0.71	0.70	0.69	0.68	0.67
0.3	0.4	0.1	0.90	0.89	0.88	0.87	0.86	0.86	0.85	0.84	0.83	0.83
0.3	0.4	0.2	0.84	0.83	0.82	0.81	0.80	0.79	0.78	0.77	0.77	0.76
0.3	0.4	0.3	0.82	0.81	0.80	0.78	0.77	0.77	0.76	0.75	0.74	0.73
0.3	0.4	0.4	0.81	0.80	0.79	0.78	0.77	0.76	0.75	0.75	0.74	0.73
0.3	0.4	0.5	0.82	0.81	0.80	0.79	0.78	0.77	0.77	0.76	0.75	0.75
0.3	0.5	0.1	0.93	0.92	0.92	0.91	0.90	0.90	0.89	0.89	0.88	0.88
0.3	0.5	0.2	0.89	0.88	0.87	0.87	0.86	0.85	0.84	0.84	0.83	0.83
0.3	0.5	0.3	0.87	0.86	0.85	0.85	0.84	0.83	0.83	0.82	0.81	0.81
0.3	0.5	0.4	0.87	0.86	0.85	0.84	0.84	0.83	0.82	0.82	0.81	0.81
0.3	0.5	0.5	0.87	0.86	0.86	0.85	0.84	0.84	0.83	0.83	0.82	0.82
0.3	0.6	0.1	0.96	0.96	0.96	0.95	0.95	0.95	0.94	0.94	0.94	0.94
0.3	0.6	0.2	0.94	0.94	0.93	0.93	0.92	0.92	0.92	0.91	0.91	0.91
0.3	0.6	0.3	0.93	0.93	0.92	0.92	0.91	0.91	0.90	0.90	0.90	0.89
0.3	0.6	0.4	0.93	0.92	0.92	0.91	0.91	0.91	0.90	0.90	0.90	0.89
0.3	0.6	0.5	0.93	0.93	0.92	0.92	0.91	0.91	0.91	0.91	0.90	0.90
0.3	0.7	0.1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.3	0.7	0.2	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.3	0.7	0.3	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.3	0.7	0.4	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.3	0.7	0.5	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.3	0.8	0.1	1.04	1.04	1.05	1.05	1.06	1.06	1.06	1.07	1.07	1.07
0.3	0.8	0.2	1.07	1.08	1.08	1.09	1.09	1.10	1.10	1.11	1.11	1.12
0.3	0.8	0.3	1.09	1.09	1.10	1.11	1.11	1.12	1.13	1.13	1.14	1.14
0.3	0.8	0.4	1.10	1.10	1.11	1.12	1.12	1.13	1.13	1.14	1.14	1.15
0.3	0.8	0.5	1.10	1.10	1.11	1.11	1.12	1.13	1.13	1.13	1.14	1.14
0.3	0.9	0.1	1.09	1.09	1.10	1.11	1.12	1.13	1.14	1.14	1.15	1.16
0.3	0.9	0.2	1.15	1.17	1.18	1.20	1.21	1.22	1.24	1.25	1.26	1.28
0.3	0.9	0.3	1.20	1.22	1.24	1.26	1.27	1.29	1.31	1.32	1.34	1.35
0.3	0.9	0.4	1.24	1.26	1.28	1.29	1.31	1.33	1.35	1.36	1.38	1.40
0.3	0.9	0.5	1.25	1.27	1.29	1.31	1.32	1.34	1.36	1.37	1.39	1.40
0.3	0.95	0.1	1.11	1.12	1.13	1.14	1.15	1.16	1.18	1.19	1.20	1.21
0.3	0.95	0.2	1.21	1.22	1.24	1.26	1.28	1.30	1.32	1.34	1.36	1.38
0.3	0.95	0.3	1.28	1.31	1.33	1.36	1.39	1.41	1.44	1.46	1.48	1.51
0.3	0.95	0.4	1.34	1.37	1.40	1.43	1.46	1.49	1.52	1.54	1.57	1.60
0.3	0.95	0.5	1.38	1.41	1.44	1.47	1.51	1.53	1.56	1.59	1.62	1.65
0.3	0.99	0.1	1.13	1.14	1.16	1.17	1.18	1.20	1.21	1.22	1.24	1.25
0.3	0.99	0.2	1.25	1.28	1.30	1.33	1.35	1.37	1.40	1.42	1.45	1.47
0.3	0.99	0.3	1.36	1.39	1.43	1.46	1.50	1.53	1.57	1.60	1.64	1.67
0.3	0.99	0.4	1.45	1.50	1.54	1.59	1.63	1.67	1.72	1.76	1.80	1.85
0.3	0.99	0.5	1.54	1.59	1.64	1.69	1.74	1.79	1.84	1.89	1.94	1.99

Table 3-6. Effect of non-differential misclassification of categorical data continued
P_N=the proportion exposed in the non-diseased group

		OR_T	3.0	3.1	3.2	3.3	3.4	3.5	3.6	3.7	3.8	3.9
Specificity	Sensitivity	P_N	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X
0.3	0.1	0.1	0.69	0.68	0.67	0.66	0.65	0.64	0.63	0.62	0.61	0.61
0.3	0.1	0.2	0.58	0.56	0.55	0.54	0.53	0.52	0.52	0.51	0.50	0.49
0.3	0.1	0.3	0.52	0.51	0.50	0.49	0.48	0.48	0.47	0.46	0.45	0.44
0.3	0.1	0.4	0.50	0.49	0.48	0.47	0.46	0.46	0.45	0.44	0.43	0.43
0.3	0.1	0.5	0.50	0.49	0.48	0.47	0.46	0.46	0.45	0.44	0.44	0.43
0.3	0.2	0.1	0.73	0.72	0.71	0.70	0.69	0.69	0.68	0.67	0.66	0.66
0.3	0.2	0.2	0.63	0.62	0.61	0.60	0.59	0.58	0.58	0.57	0.56	0.55
0.3	0.2	0.3	0.59	0.58	0.57	0.56	0.55	0.55	0.54	0.53	0.52	0.52
0.3	0.2	0.4	0.58	0.57	0.56	0.55	0.55	0.54	0.53	0.52	0.52	0.51
0.3	0.2	0.5	0.59	0.58	0.57	0.57	0.56	0.55	0.55	0.54	0.53	0.53
0.3	0.3	0.1	0.77	0.76	0.76	0.75	0.74	0.74	0.73	0.72	0.72	0.71
0.3	0.3	0.2	0.69	0.68	0.67	0.66	0.65	0.65	0.64	0.63	0.63	0.62
0.3	0.3	0.3	0.66	0.65	0.64	0.63	0.62	0.62	0.61	0.60	0.60	0.59
0.3	0.3	0.4	0.65	0.64	0.64	0.63	0.62	0.62	0.61	0.61	0.60	0.59
0.3	0.3	0.5	0.67	0.66	0.65	0.65	0.64	0.64	0.63	0.63	0.62	0.62
0.3	0.4	0.1	0.82	0.81	0.81	0.80	0.80	0.79	0.78	0.78	0.77	0.77
0.3	0.4	0.2	0.75	0.74	0.74	0.73	0.72	0.72	0.71	0.70	0.70	0.69
0.3	0.4	0.3	0.72	0.72	0.71	0.70	0.70	0.69	0.69	0.68	0.68	0.67
0.3	0.4	0.4	0.72	0.72	0.71	0.71	0.70	0.70	0.69	0.69	0.68	0.68
0.3	0.4	0.5	0.74	0.73	0.73	0.72	0.72	0.72	0.71	0.71	0.70	0.70
0.3	0.5	0.1	0.87	0.87	0.86	0.86	0.86	0.85	0.85	0.84	0.84	0.83
0.3	0.5	0.2	0.82	0.81	0.81	0.80	0.80	0.79	0.79	0.79	0.78	0.78
0.3	0.5	0.3	0.80	0.80	0.79	0.79	0.78	0.78	0.77	0.77	0.76	0.76
0.3	0.5	0.4	0.80	0.80	0.79	0.79	0.78	0.78	0.78	0.77	0.77	0.77
0.3	0.5	0.5	0.81	0.81	0.81	0.80	0.80	0.80	0.79	0.79	0.79	0.79
0.3	0.6	0.1	0.93	0.93	0.93	0.93	0.92	0.92	0.92	0.92	0.91	0.91
0.3	0.6	0.2	0.90	0.90	0.90	0.89	0.89	0.89	0.88	0.88	0.88	0.88
0.3	0.6	0.3	0.89	0.89	0.88	0.88	0.88	0.88	0.87	0.87	0.87	0.87
0.3	0.6	0.4	0.89	0.89	0.88	0.88	0.88	0.88	0.87	0.87	0.87	0.87
0.3	0.6	0.5	0.90	0.90	0.89	0.89	0.89	0.89	0.89	0.88	0.88	0.88
0.3	0.7	0.1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.3	0.7	0.2	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.3	0.7	0.3	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.3	0.7	0.4	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.3	0.7	0.5	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.3	0.8	0.1	1.08	1.08	1.08	1.09	1.09	1.09	1.10	1.10	1.10	1.11
0.3	0.8	0.2	1.12	1.13	1.13	1.14	1.14	1.15	1.15	1.15	1.16	1.16
0.3	0.8	0.3	1.15	1.15	1.16	1.16	1.17	1.17	1.18	1.18	1.18	1.19
0.3	0.8	0.4	1.15	1.16	1.16	1.17	1.17	1.18	1.18	1.18	1.19	1.19
0.3	0.8	0.5	1.15	1.15	1.16	1.16	1.16	1.17	1.17	1.17	1.18	1.18
0.3	0.9	0.1	1.17	1.17	1.18	1.19	1.20	1.20	1.21	1.22	1.23	1.23
0.3	0.9	0.2	1.29	1.30	1.31	1.33	1.34	1.35	1.36	1.37	1.38	1.39
0.3	0.9	0.3	1.37	1.38	1.40	1.41	1.43	1.44	1.45	1.46	1.48	1.49
0.3	0.9	0.4	1.41	1.43	1.44	1.45	1.47	1.48	1.49	1.51	1.52	1.53
0.3	0.9	0.5	1.42	1.43	1.44	1.46	1.47	1.48	1.49	1.50	1.51	1.53
0.3	0.95	0.1	1.22	1.23	1.24	1.25	1.26	1.27	1.28	1.29	1.30	1.31
0.3	0.95	0.2	1.40	1.41	1.43	1.45	1.47	1.48	1.50	1.52	1.54	1.55
0.3	0.95	0.3	1.53	1.55	1.58	1.60	1.62	1.65	1.67	1.69	1.71	1.73
0.3	0.95	0.4	1.63	1.65	1.68	1.70	1.73	1.75	1.77	1.80	1.82	1.84
0.3	0.95	0.5	1.67	1.70	1.72	1.75	1.77	1.80	1.82	1.84	1.87	1.89
0.3	0.99	0.1	1.26	1.28	1.29	1.30	1.31	1.33	1.34	1.35	1.37	1.38
0.3	0.99	0.2	1.50	1.52	1.55	1.57	1.60	1.62	1.64	1.67	1.69	1.72
0.3	0.99	0.3	1.71	1.74	1.78	1.81	1.84	1.88	1.91	1.94	1.98	2.01
0.3	0.99	0.4	1.89	1.93	1.97	2.02	2.06	2.10	2.14	2.18	2.22	2.26
0.3	0.99	0.5	2.04	2.09	2.14	2.19	2.23	2.28	2.33	2.37	2.42	2.47

Table 3-6. Effect of non-differential misclassification of categorical data continued

P_N=the proportion exposed in the non-diseased grou OR_X=Observed OR

		OR_T	1.05	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9
Specificity	**Sensitivity**	P_N	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X
0.4	0.1	0.1	0.99	0.98	0.97	0.95	0.93	0.92	0.90	0.89	0.87	0.86
0.4	0.1	0.2	0.98	0.97	0.94	0.91	0.89	0.86	0.84	0.82	0.80	0.78
0.4	0.1	0.3	0.98	0.96	0.92	0.89	0.86	0.83	0.80	0.78	0.76	0.74
0.4	0.1	0.4	0.98	0.95	0.91	0.87	0.84	0.81	0.78	0.75	0.73	0.71
0.4	0.1	0.5	0.97	0.95	0.90	0.86	0.83	0.80	0.77	0.74	0.72	0.70
0.4	0.2	0.1	0.99	0.99	0.97	0.96	0.95	0.93	0.92	0.91	0.90	0.89
0.4	0.2	0.2	0.99	0.98	0.95	0.93	0.91	0.89	0.87	0.85	0.84	0.82
0.4	0.2	0.3	0.98	0.97	0.94	0.91	0.89	0.86	0.84	0.82	0.80	0.79
0.4	0.2	0.4	0.98	0.96	0.93	0.90	0.87	0.85	0.83	0.81	0.79	0.77
0.4	0.2	0.5	0.98	0.96	0.93	0.90	0.87	0.84	0.82	0.80	0.78	0.77
0.4	0.3	0.1	0.99	0.99	0.98	0.97	0.96	0.95	0.94	0.93	0.92	0.91
0.4	0.3	0.2	0.99	0.98	0.96	0.95	0.93	0.92	0.90	0.89	0.88	0.86
0.4	0.3	0.3	0.99	0.98	0.95	0.93	0.91	0.90	0.88	0.86	0.85	0.84
0.4	0.3	0.4	0.99	0.97	0.95	0.93	0.91	0.89	0.87	0.85	0.84	0.83
0.4	0.3	0.5	0.99	0.97	0.95	0.92	0.90	0.89	0.87	0.85	0.84	0.83
0.4	0.4	0.1	1.00	0.99	0.99	0.98	0.97	0.97	0.96	0.95	0.95	0.94
0.4	0.4	0.2	0.99	0.99	0.98	0.96	0.95	0.94	0.93	0.92	0.91	0.91
0.4	0.4	0.3	0.99	0.98	0.97	0.95	0.94	0.93	0.92	0.91	0.90	0.89
0.4	0.4	0.4	0.99	0.98	0.97	0.95	0.94	0.92	0.91	0.90	0.89	0.88
0.4	0.4	0.5	0.99	0.98	0.96	0.95	0.94	0.92	0.91	0.90	0.89	0.88
0.4	0.5	0.1	1.00	1.00	0.99	0.99	0.99	0.98	0.98	0.98	0.97	0.97
0.4	0.5	0.2	1.00	0.99	0.99	0.98	0.98	0.97	0.97	0.96	0.96	0.95
0.4	0.5	0.3	1.00	0.99	0.98	0.98	0.97	0.96	0.96	0.95	0.95	0.94
0.4	0.5	0.4	1.00	0.99	0.98	0.97	0.97	0.96	0.95	0.95	0.94	0.94
0.4	0.5	0.5	1.00	0.99	0.98	0.97	0.97	0.96	0.95	0.95	0.94	0.94
0.4	0.6	0.1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.4	0.6	0.2	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.4	0.6	0.3	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.4	0.6	0.4	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.4	0.6	0.5	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.4	0.7	0.1	1.00	1.00	1.01	1.01	1.01	1.02	1.02	1.03	1.03	1.03
0.4	0.7	0.2	1.00	1.01	1.01	1.02	1.03	1.03	1.04	1.04	1.05	1.05
0.4	0.7	0.3	1.00	1.01	1.02	1.03	1.03	1.04	1.05	1.05	1.06	1.07
0.4	0.7	0.4	1.01	1.01	1.02	1.03	1.04	1.04	1.05	1.06	1.07	1.07
0.4	0.7	0.5	1.01	1.01	1.02	1.03	1.04	1.05	1.05	1.06	1.07	1.07
0.4	0.8	0.1	1.00	1.01	1.02	1.02	1.03	1.04	1.04	1.05	1.06	1.07
0.4	0.8	0.2	1.01	1.01	1.03	1.04	1.05	1.07	1.08	1.09	1.10	1.11
0.4	0.8	0.3	1.01	1.02	1.04	1.05	1.07	1.09	1.10	1.12	1.13	1.15
0.4	0.8	0.4	1.01	1.02	1.04	1.06	1.08	1.10	1.12	1.13	1.15	1.16
0.4	0.8	0.5	1.01	1.02	1.04	1.06	1.08	1.10	1.12	1.14	1.15	1.16
0.4	0.9	0.1	1.01	1.01	1.02	1.03	1.05	1.06	1.07	1.08	1.09	1.10
0.4	0.9	0.2	1.01	1.02	1.04	1.06	1.08	1.10	1.12	1.14	1.16	1.18
0.4	0.9	0.3	1.01	1.03	1.06	1.09	1.11	1.14	1.17	1.19	1.22	1.24
0.4	0.9	0.4	1.02	1.04	1.07	1.10	1.14	1.17	1.20	1.23	1.26	1.28
0.4	0.9	0.5	1.02	1.04	1.08	1.11	1.15	1.18	1.21	1.25	1.28	1.31
0.4	0.95	0.1	1.01	1.01	1.03	1.04	1.05	1.07	1.08	1.09	1.11	1.12
0.4	0.95	0.2	1.01	1.03	1.05	1.08	1.10	1.12	1.15	1.17	1.20	1.22
0.4	0.95	0.3	1.02	1.04	1.07	1.10	1.14	1.17	1.21	1.24	1.27	1.30
0.4	0.95	0.4	1.02	1.04	1.09	1.13	1.17	1.21	1.25	1.29	1.33	1.37
0.4	0.95	0.5	1.02	1.05	1.10	1.15	1.19	1.24	1.28	1.33	1.37	1.41
0.4	0.99	0.1	1.01	1.02	1.03	1.05	1.06	1.08	1.09	1.11	1.12	1.14
0.4	0.99	0.2	1.01	1.03	1.06	1.09	1.11	1.14	1.17	1.20	1.23	1.26
0.4	0.99	0.3	1.02	1.04	1.08	1.12	1.16	1.20	1.24	1.28	1.32	1.36
0.4	0.99	0.4	1.03	1.05	1.10	1.15	1.20	1.25	1.30	1.35	1.40	1.45
0.4	0.99	0.5	1.03	1.06	1.12	1.18	1.24	1.30	1.35	1.41	1.47	1.53

Table 3-6. Effect of non-differential misclassification of categorical data continued

P_N=the proportion exposed in the non-diseased group

		OR_T	2.0	2.1	2.2	2.3	2.4	2.5	2.6	2.7	2.8	2.9
Specificity	Sensitivity	P_N	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X
0.4	0.1	0.1	0.85	0.84	0.82	0.81	0.80	0.79	0.78	0.77	0.76	0.75
0.4	0.1	0.2	0.76	0.75	0.73	0.72	0.70	0.69	0.68	0.66	0.65	0.64
0.4	0.1	0.3	0.72	0.70	0.68	0.66	0.65	0.63	0.62	0.61	0.59	0.58
0.4	0.1	0.4	0.69	0.67	0.65	0.63	0.62	0.61	0.59	0.58	0.57	0.56
0.4	0.1	0.5	0.68	0.66	0.64	0.62	0.61	0.60	0.58	0.57	0.56	0.55
0.4	0.2	0.1	0.88	0.87	0.86	0.85	0.84	0.83	0.82	0.81	0.80	0.79
0.4	0.2	0.2	0.81	0.79	0.78	0.77	0.76	0.74	0.73	0.72	0.71	0.70
0.4	0.2	0.3	0.77	0.75	0.74	0.73	0.71	0.70	0.69	0.68	0.67	0.66
0.4	0.2	0.4	0.75	0.74	0.72	0.71	0.70	0.69	0.67	0.66	0.65	0.65
0.4	0.2	0.5	0.75	0.74	0.72	0.71	0.70	0.69	0.68	0.67	0.66	0.65
0.4	0.3	0.1	0.91	0.90	0.89	0.88	0.87	0.87	0.86	0.85	0.85	0.84
0.4	0.3	0.2	0.85	0.84	0.83	0.82	0.81	0.80	0.79	0.78	0.78	0.77
0.4	0.3	0.3	0.82	0.81	0.80	0.79	0.78	0.77	0.76	0.75	0.74	0.74
0.4	0.3	0.4	0.81	0.80	0.79	0.78	0.77	0.76	0.75	0.74	0.74	0.73
0.4	0.3	0.5	0.81	0.80	0.79	0.78	0.78	0.77	0.76	0.75	0.75	0.74
0.4	0.4	0.1	0.94	0.93	0.92	0.92	0.91	0.91	0.90	0.90	0.89	0.89
0.4	0.4	0.2	0.90	0.89	0.88	0.88	0.87	0.86	0.86	0.85	0.84	0.84
0.4	0.4	0.3	0.88	0.87	0.86	0.85	0.85	0.84	0.83	0.83	0.82	0.82
0.4	0.4	0.4	0.87	0.86	0.86	0.85	0.84	0.84	0.83	0.82	0.82	0.81
0.4	0.4	0.5	0.88	0.87	0.86	0.85	0.85	0.84	0.84	0.83	0.83	0.82
0.4	0.5	0.1	0.97	0.96	0.96	0.96	0.96	0.95	0.95	0.95	0.95	0.94
0.4	0.5	0.2	0.95	0.94	0.94	0.93	0.93	0.93	0.92	0.92	0.92	0.91
0.4	0.5	0.3	0.94	0.93	0.93	0.92	0.92	0.92	0.91	0.91	0.91	0.90
0.4	0.5	0.4	0.93	0.93	0.92	0.92	0.92	0.91	0.91	0.91	0.90	0.90
0.4	0.5	0.5	0.94	0.93	0.93	0.92	0.92	0.92	0.91	0.91	0.91	0.91
0.4	0.6	0.1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.4	0.6	0.2	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.4	0.6	0.3	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.4	0.6	0.4	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.4	0.6	0.5	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.4	0.7	0.1	1.04	1.04	1.04	1.04	1.05	1.05	1.05	1.06	1.06	1.06
0.4	0.7	0.2	1.06	1.06	1.07	1.07	1.08	1.08	1.09	1.09	1.10	1.10
0.4	0.7	0.3	1.07	1.08	1.08	1.09	1.09	1.10	1.10	1.11	1.11	1.12
0.4	0.7	0.4	1.08	1.08	1.09	1.09	1.10	1.10	1.11	1.11	1.12	1.12
0.4	0.7	0.5	1.08	1.08	1.09	1.09	1.10	1.10	1.10	1.11	1.11	1.12
0.4	0.8	0.1	1.07	1.08	1.09	1.09	1.10	1.11	1.11	1.12	1.13	1.13
0.4	0.8	0.2	1.13	1.14	1.15	1.16	1.17	1.18	1.19	1.20	1.21	1.22
0.4	0.8	0.3	1.16	1.17	1.19	1.20	1.21	1.22	1.23	1.24	1.26	1.27
0.4	0.8	0.4	1.18	1.19	1.20	1.22	1.23	1.24	1.25	1.26	1.27	1.28
0.4	0.8	0.5	1.18	1.19	1.20	1.22	1.23	1.24	1.25	1.26	1.27	1.28
0.4	0.9	0.1	1.11	1.12	1.13	1.15	1.16	1.17	1.18	1.19	1.20	1.21
0.4	0.9	0.2	1.20	1.22	1.24	1.26	1.28	1.29	1.31	1.33	1.35	1.37
0.4	0.9	0.3	1.27	1.29	1.32	1.34	1.36	1.39	1.41	1.43	1.45	1.47
0.4	0.9	0.4	1.31	1.34	1.37	1.39	1.42	1.44	1.47	1.49	1.51	1.53
0.4	0.9	0.5	1.33	1.36	1.39	1.41	1.44	1.46	1.48	1.51	1.53	1.55
0.4	0.95	0.1	1.13	1.15	1.16	1.17	1.19	1.20	1.21	1.23	1.24	1.25
0.4	0.95	0.2	1.25	1.27	1.29	1.32	1.34	1.36	1.39	1.41	1.43	1.46
0.4	0.95	0.3	1.34	1.37	1.40	1.43	1.46	1.49	1.52	1.55	1.58	1.61
0.4	0.95	0.4	1.41	1.44	1.48	1.52	1.55	1.59	1.62	1.66	1.69	1.72
0.4	0.95	0.5	1.45	1.49	1.53	1.57	1.61	1.65	1.68	1.72	1.75	1.79
0.4	0.99	0.1	1.15	1.17	1.18	1.20	1.21	1.23	1.24	1.26	1.27	1.29
0.4	0.99	0.2	1.28	1.31	1.34	1.37	1.40	1.42	1.45	1.48	1.51	1.54
0.4	0.99	0.3	1.40	1.44	1.48	1.52	1.56	1.60	1.64	1.67	1.71	1.75
0.4	0.99	0.4	1.50	1.55	1.60	1.65	1.69	1.74	1.79	1.84	1.89	1.93
0.4	0.99	0.5	1.58	1.64	1.70	1.75	1.81	1.87	1.92	1.98	2.03	2.09

Table 3-6. Effect of non-differential misclassification of categorical data continued
P_N=the proportion exposed in the non-diseased group

		OR_T	3.0	3.1	3.2	3.3	3.4	3.5	3.6	3.7	3.8	3.9
Specificity	Sensitivity	P_N	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X
0.4	0.1	0.1	0.74	0.73	0.72	0.71	0.71	0.70	0.69	0.68	0.67	0.67
0.4	0.1	0.2	0.63	0.62	0.61	0.60	0.59	0.58	0.57	0.56	0.55	0.55
0.4	0.1	0.3	0.57	0.56	0.55	0.54	0.53	0.52	0.52	0.51	0.50	0.49
0.4	0.1	0.4	0.55	0.54	0.53	0.52	0.51	0.50	0.49	0.48	0.48	0.47
0.4	0.1	0.5	0.54	0.53	0.52	0.51	0.50	0.50	0.49	0.48	0.48	0.47
0.4	0.2	0.1	0.79	0.78	0.77	0.76	0.76	0.75	0.74	0.74	0.73	0.72
0.4	0.2	0.2	0.69	0.68	0.67	0.67	0.66	0.65	0.64	0.64	0.63	0.62
0.4	0.2	0.3	0.65	0.64	0.63	0.62	0.62	0.61	0.60	0.60	0.59	0.58
0.4	0.2	0.4	0.64	0.63	0.62	0.61	0.61	0.60	0.59	0.59	0.58	0.57
0.4	0.2	0.5	0.64	0.64	0.63	0.62	0.62	0.61	0.60	0.60	0.59	0.59
0.4	0.3	0.1	0.83	0.83	0.82	0.82	0.81	0.80	0.80	0.79	0.79	0.78
0.4	0.3	0.2	0.76	0.75	0.75	0.74	0.73	0.73	0.72	0.71	0.71	0.70
0.4	0.3	0.3	0.73	0.72	0.71	0.71	0.70	0.70	0.69	0.68	0.68	0.67
0.4	0.3	0.4	0.72	0.72	0.71	0.70	0.70	0.69	0.69	0.68	0.68	0.67
0.4	0.3	0.5	0.73	0.73	0.72	0.72	0.71	0.71	0.70	0.70	0.69	0.69
0.4	0.4	0.1	0.89	0.88	0.88	0.87	0.87	0.86	0.86	0.86	0.85	0.85
0.4	0.4	0.2	0.83	0.83	0.82	0.82	0.81	0.81	0.80	0.80	0.79	0.79
0.4	0.4	0.3	0.81	0.81	0.80	0.80	0.79	0.79	0.78	0.78	0.77	0.77
0.4	0.4	0.4	0.81	0.80	0.80	0.79	0.79	0.79	0.78	0.78	0.78	0.77
0.4	0.4	0.5	0.82	0.81	0.81	0.81	0.80	0.80	0.80	0.79	0.79	0.79
0.4	0.5	0.1	0.94	0.94	0.94	0.93	0.93	0.93	0.93	0.92	0.92	0.92
0.4	0.5	0.2	0.91	0.91	0.91	0.90	0.90	0.90	0.89	0.89	0.89	0.89
0.4	0.5	0.3	0.90	0.90	0.89	0.89	0.89	0.89	0.88	0.88	0.88	0.88
0.4	0.5	0.4	0.90	0.90	0.89	0.89	0.89	0.89	0.88	0.88	0.88	0.88
0.4	0.5	0.5	0.90	0.90	0.90	0.90	0.90	0.89	0.89	0.89	0.89	0.89
0.4	0.6	0.1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.4	0.6	0.2	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.4	0.6	0.3	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.4	0.6	0.4	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.4	0.6	0.5	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.4	0.7	0.1	1.07	1.07	1.07	1.07	1.08	1.08	1.08	1.08	1.09	1.09
0.4	0.7	0.2	1.10	1.11	1.11	1.11	1.12	1.12	1.13	1.13	1.13	1.14
0.4	0.7	0.3	1.12	1.13	1.13	1.13	1.14	1.14	1.14	1.15	1.15	1.15
0.4	0.7	0.4	1.13	1.13	1.13	1.14	1.14	1.14	1.15	1.15	1.15	1.15
0.4	0.7	0.5	1.12	1.12	1.12	1.13	1.13	1.13	1.14	1.14	1.14	1.14
0.4	0.8	0.1	1.14	1.14	1.15	1.16	1.16	1.17	1.17	1.18	1.19	1.19
0.4	0.8	0.2	1.23	1.24	1.25	1.25	1.26	1.27	1.28	1.29	1.30	1.30
0.4	0.8	0.3	1.28	1.29	1.30	1.31	1.32	1.32	1.33	1.34	1.35	1.36
0.4	0.8	0.4	1.29	1.30	1.31	1.32	1.33	1.34	1.35	1.36	1.36	1.37
0.4	0.8	0.5	1.29	1.29	1.30	1.31	1.32	1.32	1.33	1.34	1.34	1.35
0.4	0.9	0.1	1.22	1.23	1.24	1.25	1.26	1.27	1.28	1.29	1.30	1.31
0.4	0.9	0.2	1.38	1.40	1.42	1.43	1.45	1.47	1.48	1.50	1.51	1.53
0.4	0.9	0.3	1.49	1.51	1.53	1.55	1.57	1.59	1.61	1.63	1.65	1.67
0.4	0.9	0.4	1.56	1.58	1.60	1.62	1.64	1.66	1.68	1.70	1.71	1.73
0.4	0.9	0.5	1.57	1.59	1.61	1.63	1.65	1.67	1.68	1.70	1.72	1.73
0.4	0.95	0.1	1.26	1.28	1.29	1.30	1.32	1.33	1.34	1.35	1.37	1.38
0.4	0.95	0.2	1.48	1.50	1.52	1.54	1.57	1.59	1.61	1.63	1.65	1.68
0.4	0.95	0.3	1.64	1.67	1.70	1.73	1.76	1.78	1.81	1.84	1.87	1.89
0.4	0.95	0.4	1.76	1.79	1.82	1.85	1.88	1.92	1.95	1.98	2.01	2.04
0.4	0.95	0.5	1.82	1.85	1.89	1.92	1.95	1.98	2.01	2.04	2.07	2.10
0.4	0.99	0.1	1.30	1.32	1.33	1.35	1.36	1.38	1.39	1.41	1.42	1.44
0.4	0.99	0.2	1.56	1.59	1.62	1.65	1.68	1.70	1.73	1.76	1.79	1.81
0.4	0.99	0.3	1.79	1.83	1.87	1.91	1.94	1.98	2.02	2.06	2.10	2.14
0.4	0.99	0.4	1.98	2.03	2.08	2.12	2.17	2.22	2.27	2.31	2.36	2.40
0.4	0.99	0.5	2.14	2.20	2.25	2.30	2.36	2.41	2.46	2.52	2.57	2.62

Table 3-6. Effect of non-differential misclassification of categorical data continued

P_N=the proportion exposed in the non-diseased group OR_X=Observed OR

		OR_T	1.05	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9
Specificity	**Sensitivity**	P_N	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X
0.5	0.1	0.1	0.99	0.99	0.97	0.96	0.95	0.93	0.92	0.91	0.90	0.89
0.5	0.1	0.2	0.99	0.97	0.95	0.93	0.91	0.89	0.87	0.85	0.83	0.82
0.5	0.1	0.3	0.98	0.97	0.93	0.91	0.88	0.85	0.83	0.81	0.79	0.77
0.5	0.1	0.4	0.98	0.96	0.92	0.89	0.86	0.83	0.81	0.78	0.76	0.74
0.5	0.1	0.5	0.98	0.96	0.92	0.88	0.85	0.82	0.79	0.77	0.75	0.73
0.5	0.2	0.1	0.99	0.99	0.98	0.97	0.96	0.95	0.94	0.93	0.92	0.91
0.5	0.2	0.2	0.99	0.98	0.96	0.95	0.93	0.91	0.90	0.89	0.87	0.86
0.5	0.2	0.3	0.99	0.97	0.95	0.93	0.91	0.89	0.87	0.86	0.84	0.83
0.5	0.2	0.4	0.99	0.97	0.94	0.92	0.90	0.88	0.86	0.84	0.83	0.81
0.5	0.2	0.5	0.98	0.97	0.94	0.92	0.89	0.87	0.86	0.84	0.82	0.81
0.5	0.3	0.1	1.00	0.99	0.99	0.98	0.97	0.97	0.96	0.95	0.95	0.94
0.5	0.3	0.2	0.99	0.99	0.98	0.96	0.95	0.94	0.93	0.92	0.91	0.91
0.5	0.3	0.3	0.99	0.98	0.97	0.95	0.94	0.93	0.92	0.91	0.90	0.89
0.5	0.3	0.4	0.99	0.98	0.96	0.95	0.93	0.92	0.91	0.90	0.89	0.88
0.5	0.3	0.5	0.99	0.98	0.96	0.95	0.93	0.92	0.91	0.90	0.89	0.88
0.5	0.4	0.1	1.00	1.00	0.99	0.99	0.99	0.98	0.98	0.98	0.97	0.97
0.5	0.4	0.2	1.00	0.99	0.99	0.98	0.98	0.97	0.97	0.96	0.96	0.95
0.5	0.4	0.3	1.00	0.99	0.98	0.98	0.97	0.96	0.96	0.95	0.95	0.94
0.5	0.4	0.4	1.00	0.99	0.98	0.97	0.97	0.96	0.95	0.95	0.94	0.94
0.5	0.4	0.5	1.00	0.99	0.98	0.97	0.97	0.96	0.95	0.95	0.94	0.94
0.5	0.5	0.1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.5	0.5	0.2	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.5	0.5	0.3	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.5	0.5	0.4	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.5	0.5	0.5	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.5	0.6	0.1	1.00	1.00	1.01	1.01	1.01	1.02	1.02	1.02	1.03	1.03
0.5	0.6	0.2	1.00	1.01	1.01	1.02	1.02	1.03	1.03	1.04	1.05	1.05
0.5	0.6	0.3	1.00	1.01	1.02	1.02	1.03	1.04	1.04	1.05	1.06	1.06
0.5	0.6	0.4	1.00	1.01	1.02	1.03	1.03	1.04	1.05	1.05	1.06	1.07
0.5	0.6	0.5	1.00	1.01	1.02	1.03	1.03	1.04	1.05	1.05	1.06	1.06
0.5	0.7	0.1	1.00	1.01	1.01	1.02	1.03	1.03	1.04	1.05	1.05	1.06
0.5	0.7	0.2	1.01	1.01	1.03	1.04	1.05	1.06	1.07	1.08	1.09	1.10
0.5	0.7	0.3	1.01	1.02	1.03	1.05	1.06	1.08	1.09	1.10	1.12	1.13
0.5	0.7	0.4	1.01	1.02	1.04	1.05	1.07	1.09	1.10	1.11	1.13	1.14
0.5	0.7	0.5	1.01	1.02	1.04	1.06	1.07	1.09	1.10	1.12	1.13	1.14
0.5	0.8	0.1	1.01	1.01	1.02	1.03	1.04	1.05	1.06	1.07	1.08	1.09
0.5	0.8	0.2	1.01	1.02	1.04	1.06	1.08	1.09	1.11	1.13	1.15	1.16
0.5	0.8	0.3	1.01	1.03	1.05	1.07	1.10	1.12	1.14	1.17	1.19	1.21
0.5	0.8	0.4	1.02	1.03	1.06	1.09	1.11	1.14	1.16	1.19	1.21	1.23
0.5	0.8	0.5	1.02	1.03	1.06	1.09	1.12	1.14	1.17	1.19	1.21	1.24
0.5	0.9	0.1	1.01	1.01	1.03	1.04	1.06	1.07	1.09	1.10	1.11	1.13
0.5	0.9	0.2	1.01	1.03	1.05	1.08	1.10	1.13	1.15	1.18	1.20	1.23
0.5	0.9	0.3	1.02	1.04	1.07	1.10	1.14	1.17	1.20	1.24	1.27	1.30
0.5	0.9	0.4	1.02	1.04	1.08	1.12	1.16	1.20	1.24	1.28	1.31	1.35
0.5	0.9	0.5	1.02	1.05	1.09	1.14	1.18	1.22	1.26	1.30	1.34	1.37
0.5	0.95	0.1	1.01	1.02	1.03	1.05	1.07	1.08	1.10	1.11	1.13	1.15
0.5	0.95	0.2	1.01	1.03	1.06	1.09	1.12	1.15	1.18	1.20	1.23	1.26
0.5	0.95	0.3	1.02	1.04	1.08	1.12	1.16	1.20	1.24	1.28	1.32	1.35
0.5	0.95	0.4	1.02	1.05	1.10	1.15	1.19	1.24	1.29	1.33	1.38	1.42
0.5	0.95	0.5	1.03	1.06	1.11	1.16	1.22	1.27	1.32	1.37	1.42	1.47
0.5	0.99	0.1	1.01	1.02	1.04	1.05	1.07	1.09	1.11	1.12	1.14	1.16
0.5	0.99	0.2	1.02	1.03	1.07	1.10	1.13	1.16	1.20	1.23	1.26	1.29
0.5	0.99	0.3	1.02	1.05	1.09	1.13	1.18	1.22	1.27	1.31	1.36	1.40
0.5	0.99	0.4	1.03	1.06	1.11	1.17	1.22	1.28	1.33	1.39	1.44	1.49
0.5	0.99	0.5	1.03	1.06	1.13	1.19	1.26	1.32	1.38	1.45	1.51	1.57

Table 3-6. Effect of non-differential misclassification of categorical data continued

P_N=the proportion exposed in the non-diseased group

		OR_T	2.0	2.1	2.2	2.3	2.4	2.5	2.6	2.7	2.8	2.9
Specificity	**Sensitivity**	P_N	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X
0.5	0.1	0.1	0.88	0.87	0.86	0.85	0.84	0.83	0.82	0.81	0.80	0.79
0.5	0.1	0.2	0.80	0.78	0.77	0.76	0.74	0.73	0.72	0.71	0.70	0.69
0.5	0.1	0.3	0.75	0.73	0.72	0.70	0.69	0.68	0.66	0.65	0.64	0.63
0.5	0.1	0.4	0.72	0.71	0.69	0.67	0.66	0.65	0.63	0.62	0.61	0.60
0.5	0.1	0.5	0.71	0.69	0.68	0.66	0.65	0.64	0.62	0.61	0.60	0.59
0.5	0.2	0.1	0.91	0.90	0.89	0.88	0.87	0.87	0.86	0.85	0.85	0.84
0.5	0.2	0.2	0.85	0.84	0.83	0.82	0.81	0.80	0.79	0.78	0.77	0.76
0.5	0.2	0.3	0.81	0.80	0.79	0.78	0.77	0.76	0.75	0.74	0.73	0.72
0.5	0.2	0.4	0.80	0.79	0.77	0.76	0.75	0.74	0.73	0.72	0.71	0.71
0.5	0.2	0.5	0.80	0.78	0.77	0.76	0.75	0.74	0.73	0.73	0.72	0.71
0.5	0.3	0.1	0.94	0.93	0.93	0.92	0.92	0.91	0.91	0.90	0.90	0.89
0.5	0.3	0.2	0.90	0.89	0.88	0.87	0.87	0.86	0.85	0.85	0.84	0.84
0.5	0.3	0.3	0.88	0.87	0.86	0.85	0.84	0.84	0.83	0.82	0.82	0.81
0.5	0.3	0.4	0.87	0.86	0.85	0.84	0.84	0.83	0.82	0.82	0.81	0.80
0.5	0.3	0.5	0.87	0.86	0.85	0.85	0.84	0.83	0.83	0.82	0.82	0.81
0.5	0.4	0.1	0.97	0.96	0.96	0.96	0.96	0.95	0.95	0.95	0.95	0.94
0.5	0.4	0.2	0.95	0.94	0.94	0.94	0.93	0.93	0.93	0.92	0.92	0.92
0.5	0.4	0.3	0.94	0.93	0.93	0.92	0.92	0.92	0.91	0.91	0.91	0.90
0.5	0.4	0.4	0.93	0.93	0.92	0.92	0.92	0.91	0.91	0.91	0.90	0.90
0.5	0.4	0.5	0.93	0.93	0.93	0.92	0.92	0.92	0.91	0.91	0.91	0.91
0.5	0.5	0.1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.5	0.5	0.2	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.5	0.5	0.3	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.5	0.5	0.4	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.5	0.5	0.5	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.5	0.6	0.1	1.03	1.04	1.04	1.04	1.05	1.05	1.05	1.05	1.06	1.06
0.5	0.6	0.2	1.05	1.06	1.06	1.07	1.07	1.08	1.08	1.08	1.09	1.09
0.5	0.6	0.3	1.07	1.07	1.08	1.08	1.09	1.09	1.10	1.10	1.10	1.11
0.5	0.6	0.4	1.07	1.08	1.08	1.09	1.09	1.10	1.10	1.10	1.11	1.11
0.5	0.6	0.5	1.07	1.07	1.08	1.08	1.09	1.09	1.09	1.10	1.10	1.10
0.5	0.7	0.1	1.07	1.07	1.08	1.09	1.09	1.10	1.10	1.11	1.12	1.12
0.5	0.7	0.2	1.11	1.12	1.13	1.14	1.15	1.16	1.17	1.18	1.19	1.20
0.5	0.7	0.3	1.14	1.15	1.16	1.18	1.19	1.20	1.21	1.22	1.22	1.23
0.5	0.7	0.4	1.15	1.16	1.18	1.19	1.20	1.21	1.22	1.23	1.23	1.24
0.5	0.7	0.5	1.15	1.16	1.17	1.18	1.19	1.20	1.21	1.22	1.22	1.23
0.5	0.8	0.1	1.10	1.11	1.12	1.13	1.14	1.15	1.16	1.17	1.18	1.19
0.5	0.8	0.2	1.18	1.19	1.21	1.23	1.24	1.26	1.27	1.29	1.30	1.32
0.5	0.8	0.3	1.23	1.25	1.27	1.28	1.30	1.32	1.34	1.35	1.37	1.39
0.5	0.8	0.4	1.25	1.27	1.29	1.31	1.33	1.35	1.37	1.38	1.40	1.41
0.5	0.8	0.5	1.26	1.28	1.29	1.31	1.33	1.35	1.36	1.38	1.39	1.41
0.5	0.9	0.1	1.14	1.16	1.17	1.18	1.20	1.21	1.22	1.24	1.25	1.26
0.5	0.9	0.2	1.25	1.27	1.30	1.32	1.34	1.37	1.39	1.41	1.44	1.46
0.5	0.9	0.3	1.33	1.36	1.39	1.42	1.45	1.48	1.51	1.53	1.56	1.59
0.5	0.9	0.4	1.38	1.42	1.45	1.48	1.51	1.55	1.58	1.61	1.64	1.66
0.5	0.9	0.5	1.41	1.44	1.48	1.51	1.54	1.57	1.60	1.63	1.66	1.69
0.5	0.95	0.1	1.16	1.18	1.19	1.21	1.23	1.24	1.26	1.27	1.29	1.30
0.5	0.95	0.2	1.29	1.32	1.35	1.38	1.40	1.43	1.46	1.49	1.51	1.54
0.5	0.95	0.3	1.39	1.43	1.47	1.50	1.54	1.58	1.61	1.65	1.68	1.72
0.5	0.95	0.4	1.47	1.51	1.55	1.60	1.64	1.68	1.72	1.76	1.80	1.84
0.5	0.95	0.5	1.52	1.56	1.61	1.66	1.70	1.74	1.79	1.83	1.87	1.91
0.5	0.99	0.1	1.18	1.20	1.21	1.23	1.25	1.27	1.28	1.30	1.32	1.34
0.5	0.99	0.2	1.32	1.36	1.39	1.42	1.45	1.49	1.52	1.55	1.58	1.61
0.5	0.99	0.3	1.45	1.49	1.54	1.58	1.62	1.67	1.71	1.76	1.80	1.84
0.5	0.99	0.4	1.55	1.60	1.66	1.71	1.76	1.82	1.87	1.92	1.98	2.03
0.5	0.99	0.5	1.63	1.69	1.76	1.82	1.88	1.94	2.00	2.06	2.12	2.18

Table 3-6. Effect of non-differential misclassification of categorical data continued

P_N=the proportion exposed in the non-diseased group

		OR_T	3.0	3.1	3.2	3.3	3.4	3.5	3.6	3.7	3.8	3.9
Specificity	**Sensitivity**	P_N	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X
0.5	0.1	0.1	0.78	0.77	0.77	0.76	0.75	0.74	0.74	0.73	0.72	0.72
0.5	0.1	0.2	0.68	0.67	0.66	0.65	0.64	0.63	0.62	0.61	0.61	0.60
0.5	0.1	0.3	0.62	0.61	0.60	0.59	0.58	0.57	0.57	0.56	0.55	0.54
0.5	0.1	0.4	0.59	0.58	0.57	0.56	0.56	0.55	0.54	0.53	0.53	0.52
0.5	0.1	0.5	0.58	0.57	0.57	0.56	0.55	0.54	0.54	0.53	0.52	0.52
0.5	0.2	0.1	0.83	0.83	0.82	0.81	0.81	0.80	0.80	0.79	0.79	0.78
0.5	0.2	0.2	0.75	0.74	0.74	0.73	0.72	0.72	0.71	0.70	0.70	0.69
0.5	0.2	0.3	0.71	0.71	0.70	0.69	0.68	0.68	0.67	0.66	0.66	0.65
0.5	0.2	0.4	0.70	0.69	0.69	0.68	0.67	0.67	0.66	0.66	0.65	0.65
0.5	0.2	0.5	0.70	0.70	0.69	0.69	0.68	0.68	0.67	0.67	0.66	0.66
0.5	0.3	0.1	0.89	0.88	0.88	0.87	0.87	0.87	0.86	0.86	0.85	0.85
0.5	0.3	0.2	0.83	0.82	0.82	0.81	0.81	0.80	0.80	0.80	0.79	0.79
0.5	0.3	0.3	0.81	0.80	0.79	0.79	0.78	0.78	0.78	0.77	0.77	0.76
0.5	0.3	0.4	0.80	0.79	0.79	0.79	0.78	0.78	0.77	0.77	0.77	0.76
0.5	0.3	0.5	0.81	0.80	0.80	0.80	0.79	0.79	0.78	0.78	0.78	0.78
0.5	0.4	0.1	0.94	0.94	0.94	0.93	0.93	0.93	0.93	0.93	0.92	0.92
0.5	0.4	0.2	0.91	0.91	0.91	0.90	0.90	0.90	0.90	0.89	0.89	0.89
0.5	0.4	0.3	0.90	0.90	0.89	0.89	0.89	0.89	0.88	0.88	0.88	0.88
0.5	0.4	0.4	0.90	0.90	0.89	0.89	0.89	0.89	0.88	0.88	0.88	0.88
0.5	0.4	0.5	0.90	0.90	0.90	0.90	0.90	0.89	0.89	0.89	0.89	0.89
0.5	0.5	0.1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.5	0.5	0.2	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.5	0.5	0.3	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.5	0.5	0.4	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.5	0.5	0.5	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.5	0.6	0.1	1.06	1.06	1.07	1.07	1.07	1.07	1.08	1.08	1.08	1.08
0.5	0.6	0.2	1.10	1.10	1.10	1.11	1.11	1.11	1.12	1.12	1.12	1.13
0.5	0.6	0.3	1.11	1.12	1.12	1.12	1.13	1.13	1.13	1.13	1.14	1.14
0.5	0.6	0.4	1.11	1.12	1.12	1.12	1.13	1.13	1.13	1.13	1.14	1.14
0.5	0.6	0.5	1.11	1.11	1.11	1.11	1.12	1.12	1.12	1.12	1.13	1.13
0.5	0.7	0.1	1.13	1.13	1.14	1.15	1.15	1.16	1.16	1.17	1.17	1.18
0.5	0.7	0.2	1.20	1.21	1.22	1.23	1.24	1.24	1.25	1.26	1.26	1.27
0.5	0.7	0.3	1.24	1.25	1.26	1.27	1.27	1.28	1.29	1.30	1.30	1.31
0.5	0.7	0.4	1.25	1.26	1.27	1.27	1.28	1.29	1.29	1.30	1.31	1.31
0.5	0.7	0.5	1.24	1.24	1.25	1.26	1.26	1.27	1.27	1.28	1.28	1.29
0.5	0.8	0.1	1.20	1.21	1.22	1.23	1.24	1.24	1.25	1.26	1.27	1.28
0.5	0.8	0.2	1.33	1.34	1.36	1.37	1.38	1.40	1.41	1.42	1.43	1.45
0.5	0.8	0.3	1.40	1.42	1.43	1.45	1.46	1.48	1.49	1.50	1.52	1.53
0.5	0.8	0.4	1.43	1.45	1.46	1.47	1.49	1.50	1.51	1.53	1.54	1.55
0.5	0.8	0.5	1.42	1.43	1.45	1.46	1.47	1.48	1.49	1.50	1.51	1.52
0.5	0.9	0.1	1.28	1.29	1.30	1.32	1.33	1.34	1.36	1.37	1.38	1.40
0.5	0.9	0.2	1.48	1.50	1.52	1.54	1.57	1.59	1.61	1.63	1.65	1.67
0.5	0.9	0.3	1.62	1.64	1.67	1.69	1.72	1.74	1.77	1.79	1.82	1.84
0.5	0.9	0.4	1.69	1.72	1.75	1.77	1.80	1.83	1.85	1.88	1.90	1.93
0.5	0.9	0.5	1.71	1.74	1.77	1.79	1.82	1.84	1.86	1.89	1.91	1.93
0.5	0.95	0.1	1.32	1.34	1.35	1.37	1.38	1.40	1.41	1.43	1.44	1.46
0.5	0.95	0.2	1.57	1.59	1.62	1.65	1.67	1.70	1.73	1.75	1.78	1.81
0.5	0.95	0.3	1.75	1.79	1.82	1.86	1.89	1.92	1.96	1.99	2.02	2.06
0.5	0.95	0.4	1.88	1.92	1.96	2.00	2.04	2.07	2.11	2.15	2.18	2.22
0.5	0.95	0.5	1.95	1.99	2.03	2.07	2.11	2.15	2.19	2.22	2.26	2.29
0.5	0.99	0.1	1.35	1.37	1.39	1.41	1.43	1.44	1.46	1.48	1.50	1.51
0.5	0.99	0.2	1.65	1.68	1.71	1.74	1.77	1.81	1.84	1.87	1.90	1.93
0.5	0.99	0.3	1.89	1.93	1.97	2.02	2.06	2.10	2.15	2.19	2.23	2.28
0.5	0.99	0.4	2.08	2.14	2.19	2.24	2.29	2.35	2.40	2.45	2.50	2.55
0.5	0.99	0.5	2.24	2.30	2.36	2.42	2.48	2.54	2.60	2.65	2.71	2.77

Table 3-6. Effect of non-differential misclassification of categorical data continued

P_N=the proportion exposed in the non-diseased grou OR_X=Observed OR

		OR_T	1.05	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9
Specificity	**Sensitivity**	P_N	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X
0.6	0.1	0.1	0.99	0.99	0.98	0.97	0.96	0.95	0.94	0.93	0.92	0.91
0.6	0.1	0.2	0.99	0.98	0.96	0.94	0.92	0.91	0.89	0.87	0.86	0.85
0.6	0.1	0.3	0.99	0.97	0.95	0.92	0.90	0.88	0.86	0.84	0.82	0.80
0.6	0.1	0.4	0.98	0.97	0.94	0.91	0.88	0.86	0.84	0.81	0.80	0.78
0.6	0.1	0.5	0.98	0.96	0.93	0.90	0.87	0.85	0.82	0.80	0.78	0.77
0.6	0.2	0.1	1.00	0.99	0.99	0.98	0.97	0.96	0.96	0.95	0.94	0.94
0.6	0.2	0.2	0.99	0.99	0.97	0.96	0.95	0.94	0.93	0.92	0.91	0.90
0.6	0.2	0.3	0.99	0.98	0.97	0.95	0.93	0.92	0.91	0.90	0.88	0.87
0.6	0.2	0.4	0.99	0.98	0.96	0.94	0.93	0.91	0.90	0.88	0.87	0.86
0.6	0.2	0.5	0.99	0.98	0.96	0.94	0.92	0.91	0.89	0.88	0.87	0.86
0.6	0.3	0.1	1.00	1.00	0.99	0.99	0.99	0.98	0.98	0.98	0.97	0.97
0.6	0.3	0.2	1.00	0.99	0.99	0.98	0.98	0.97	0.96	0.96	0.95	0.95
0.6	0.3	0.3	1.00	0.99	0.98	0.98	0.97	0.96	0.95	0.95	0.94	0.94
0.6	0.3	0.4	0.99	0.99	0.98	0.97	0.96	0.96	0.95	0.94	0.94	0.93
0.6	0.3	0.5	0.99	0.99	0.98	0.97	0.96	0.96	0.95	0.94	0.94	0.93
0.6	0.4	0.1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.6	0.4	0.2	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.6	0.4	0.3	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.6	0.4	0.4	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.6	0.4	0.5	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.6	0.5	0.1	1.00	1.00	1.01	1.01	1.01	1.02	1.02	1.02	1.03	1.03
0.6	0.5	0.2	1.00	1.01	1.01	1.02	1.02	1.03	1.04	1.04	1.05	1.05
0.6	0.5	0.3	1.00	1.01	1.02	1.02	1.03	1.04	1.04	1.05	1.06	1.06
0.6	0.5	0.4	1.00	1.01	1.02	1.03	1.03	1.04	1.05	1.05	1.06	1.07
0.6	0.5	0.5	1.00	1.01	1.02	1.03	1.03	1.04	1.05	1.05	1.06	1.06
0.6	0.6	0.1	1.00	1.01	1.01	1.02	1.03	1.04	1.04	1.05	1.06	1.06
0.6	0.6	0.2	1.01	1.01	1.03	1.04	1.05	1.06	1.07	1.08	1.09	1.10
0.6	0.6	0.3	1.01	1.02	1.03	1.05	1.06	1.08	1.09	1.10	1.11	1.13
0.6	0.6	0.4	1.01	1.02	1.04	1.05	1.07	1.08	1.10	1.11	1.12	1.14
0.6	0.6	0.5	1.01	1.02	1.04	1.05	1.07	1.08	1.10	1.11	1.12	1.13
0.6	0.7	0.1	1.01	1.01	1.02	1.03	1.04	1.05	1.06	1.07	1.08	1.09
0.6	0.7	0.2	1.01	1.02	1.04	1.06	1.07	1.09	1.11	1.13	1.14	1.16
0.6	0.7	0.3	1.01	1.02	1.05	1.07	1.09	1.12	1.14	1.16	1.18	1.20
0.6	0.7	0.4	1.01	1.03	1.05	1.08	1.10	1.13	1.15	1.17	1.19	1.21
0.6	0.7	0.5	1.01	1.03	1.06	1.08	1.11	1.13	1.15	1.17	1.19	1.21
0.6	0.8	0.1	1.01	1.01	1.03	1.04	1.06	1.07	1.09	1.10	1.11	1.13
0.6	0.8	0.2	1.01	1.03	1.05	1.08	1.10	1.12	1.15	1.17	1.19	1.22
0.6	0.8	0.3	1.02	1.03	1.07	1.10	1.13	1.16	1.19	1.22	1.24	1.27
0.6	0.8	0.4	1.02	1.04	1.08	1.11	1.15	1.18	1.21	1.24	1.27	1.30
0.6	0.8	0.5	1.02	1.04	1.08	1.12	1.15	1.19	1.22	1.25	1.28	1.31
0.6	0.9	0.1	1.01	1.02	1.04	1.05	1.07	1.09	1.11	1.13	1.14	1.16
0.6	0.9	0.2	1.02	1.03	1.06	1.09	1.13	1.16	1.19	1.22	1.25	1.28
0.6	0.9	0.3	1.02	1.04	1.08	1.12	1.17	1.21	1.24	1.28	1.32	1.36
0.6	0.9	0.4	1.02	1.05	1.10	1.15	1.19	1.24	1.28	1.33	1.37	1.41
0.6	0.9	0.5	1.03	1.05	1.11	1.16	1.21	1.26	1.30	1.35	1.39	1.44
0.6	0.95	0.1	1.01	1.02	1.04	1.06	1.08	1.10	1.12	1.14	1.16	1.18
0.6	0.95	0.2	1.02	1.04	1.07	1.10	1.14	1.17	1.21	1.24	1.28	1.31
0.6	0.95	0.3	1.02	1.05	1.09	1.14	1.19	1.23	1.28	1.32	1.37	1.41
0.6	0.95	0.4	1.03	1.06	1.11	1.17	1.22	1.27	1.33	1.38	1.43	1.48
0.6	0.95	0.5	1.03	1.06	1.12	1.18	1.24	1.30	1.36	1.42	1.47	1.53
0.6	0.99	0.1	1.01	1.02	1.04	1.06	1.09	1.11	1.13	1.15	1.17	1.19
0.6	0.99	0.2	1.02	1.04	1.08	1.11	1.15	1.19	1.23	1.26	1.30	1.34
0.6	0.99	0.3	1.03	1.05	1.10	1.15	1.20	1.25	1.30	1.35	1.40	1.45
0.6	0.99	0.4	1.03	1.06	1.12	1.18	1.24	1.30	1.36	1.42	1.49	1.55
0.6	0.99	0.5	1.03	1.07	1.14	1.21	1.28	1.35	1.41	1.48	1.55	1.62

Table 3-6. Effect of non-differential misclassification of categorical data continued
P_N=the proportion exposed in the non-diseased group

		OR_T	2.0	2.1	2.2	2.3	2.4	2.5	2.6	2.7	2.8	2.9
Specificity	**Sensitivity**	P_N	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X
0.6	0.1	0.1	0.90	0.89	0.88	0.87	0.86	0.86	0.85	0.84	0.83	0.83
0.6	0.1	0.2	0.83	0.82	0.81	0.79	0.78	0.77	0.76	0.75	0.74	0.73
0.6	0.1	0.3	0.79	0.77	0.76	0.75	0.73	0.72	0.71	0.70	0.69	0.68
0.6	0.1	0.4	0.76	0.75	0.73	0.72	0.71	0.69	0.68	0.67	0.66	0.65
0.6	0.1	0.5	0.75	0.73	0.72	0.71	0.70	0.68	0.67	0.66	0.65	0.64
0.6	0.2	0.1	0.93	0.93	0.92	0.91	0.91	0.90	0.90	0.89	0.89	0.88
0.6	0.2	0.2	0.89	0.88	0.87	0.86	0.86	0.85	0.84	0.83	0.83	0.82
0.6	0.2	0.3	0.86	0.85	0.84	0.83	0.83	0.82	0.81	0.80	0.80	0.79
0.6	0.2	0.4	0.85	0.84	0.83	0.82	0.81	0.81	0.80	0.79	0.79	0.78
0.6	0.2	0.5	0.85	0.84	0.83	0.82	0.81	0.81	0.80	0.79	0.79	0.78
0.6	0.3	0.1	0.97	0.96	0.96	0.96	0.95	0.95	0.95	0.95	0.94	0.94
0.6	0.3	0.2	0.94	0.94	0.94	0.93	0.93	0.92	0.92	0.92	0.91	0.91
0.6	0.3	0.3	0.93	0.93	0.92	0.92	0.91	0.91	0.91	0.90	0.90	0.90
0.6	0.3	0.4	0.93	0.92	0.92	0.91	0.91	0.91	0.90	0.90	0.90	0.89
0.6	0.3	0.5	0.93	0.92	0.92	0.92	0.91	0.91	0.91	0.90	0.90	0.90
0.6	0.4	0.1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.6	0.4	0.2	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.6	0.4	0.3	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.6	0.4	0.4	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.6	0.4	0.5	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.6	0.5	0.1	1.03	1.04	1.04	1.04	1.05	1.05	1.05	1.06	1.06	1.06
0.6	0.5	0.2	1.06	1.06	1.07	1.07	1.07	1.08	1.08	1.09	1.09	1.09
0.6	0.5	0.3	1.07	1.07	1.08	1.08	1.09	1.09	1.10	1.10	1.10	1.11
0.6	0.5	0.4	1.07	1.08	1.08	1.09	1.09	1.10	1.10	1.10	1.11	1.11
0.6	0.5	0.5	1.07	1.07	1.08	1.08	1.09	1.09	1.09	1.10	1.10	1.10
0.6	0.6	0.1	1.07	1.08	1.08	1.09	1.09	1.10	1.11	1.11	1.12	1.12
0.6	0.6	0.2	1.11	1.12	1.13	1.14	1.15	1.16	1.17	1.18	1.19	1.19
0.6	0.6	0.3	1.14	1.15	1.16	1.17	1.18	1.19	1.20	1.21	1.22	1.23
0.6	0.6	0.4	1.15	1.16	1.17	1.18	1.19	1.20	1.21	1.21	1.22	1.23
0.6	0.6	0.5	1.14	1.15	1.16	1.17	1.18	1.19	1.20	1.20	1.21	1.22
0.6	0.7	0.1	1.10	1.11	1.12	1.13	1.14	1.15	1.16	1.17	1.18	1.19
0.6	0.7	0.2	1.17	1.19	1.20	1.22	1.23	1.25	1.26	1.28	1.29	1.30
0.6	0.7	0.3	1.21	1.23	1.25	1.27	1.28	1.30	1.31	1.33	1.34	1.36
0.6	0.7	0.4	1.23	1.25	1.27	1.28	1.30	1.31	1.33	1.34	1.36	1.37
0.6	0.7	0.5	1.23	1.24	1.26	1.27	1.29	1.30	1.32	1.33	1.34	1.35
0.6	0.8	0.1	1.14	1.15	1.17	1.18	1.19	1.21	1.22	1.23	1.25	1.26
0.6	0.8	0.2	1.24	1.26	1.28	1.30	1.32	1.34	1.37	1.39	1.41	1.43
0.6	0.8	0.3	1.30	1.33	1.35	1.38	1.40	1.43	1.45	1.47	1.49	1.52
0.6	0.8	0.4	1.33	1.36	1.38	1.41	1.43	1.46	1.48	1.51	1.53	1.55
0.6	0.8	0.5	1.33	1.36	1.38	1.41	1.43	1.45	1.48	1.50	1.52	1.54
0.6	0.9	0.1	1.18	1.20	1.21	1.23	1.25	1.27	1.28	1.30	1.32	1.33
0.6	0.9	0.2	1.31	1.34	1.37	1.40	1.42	1.45	1.48	1.51	1.54	1.57
0.6	0.9	0.3	1.40	1.43	1.47	1.51	1.54	1.58	1.61	1.65	1.68	1.72
0.6	0.9	0.4	1.45	1.50	1.54	1.58	1.61	1.65	1.69	1.73	1.76	1.80
0.6	0.9	0.5	1.48	1.52	1.56	1.60	1.64	1.68	1.72	1.75	1.79	1.82
0.6	0.95	0.1	1.20	1.22	1.24	1.26	1.28	1.30	1.31	1.33	1.35	1.37
0.6	0.95	0.2	1.35	1.38	1.41	1.45	1.48	1.51	1.55	1.58	1.61	1.64
0.6	0.95	0.3	1.45	1.50	1.54	1.58	1.63	1.67	1.71	1.75	1.80	1.84
0.6	0.95	0.4	1.53	1.58	1.63	1.68	1.73	1.78	1.83	1.87	1.92	1.97
0.6	0.95	0.5	1.58	1.64	1.69	1.74	1.79	1.84	1.89	1.94	1.99	2.04
0.6	0.99	0.1	1.21	1.23	1.26	1.28	1.30	1.32	1.34	1.36	1.38	1.40
0.6	0.99	0.2	1.38	1.41	1.45	1.49	1.53	1.56	1.60	1.64	1.68	1.71
0.6	0.99	0.3	1.50	1.55	1.60	1.65	1.70	1.75	1.80	1.85	1.90	1.95
0.6	0.99	0.4	1.61	1.66	1.72	1.78	1.84	1.90	1.96	2.02	2.08	2.14
0.6	0.99	0.5	1.68	1.75	1.82	1.89	1.95	2.02	2.08	2.15	2.22	2.28

Table 3-6. Effect of non-differential misclassification of categorical data continued
P_N=the proportion exposed in the non-diseased group

	OR_T		3.0	3.1	3.2	3.3	3.4	3.5	3.6	3.7	3.8	3.9
Specificity	**Sensitivity**	P_N	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X
0.6	0.1	0.1	0.82	0.81	0.81	0.80	0.79	0.79	0.78	0.77	0.77	0.76
0.6	0.1	0.2	0.72	0.71	0.71	0.70	0.69	0.68	0.67	0.67	0.66	0.65
0.6	0.1	0.3	0.67	0.66	0.65	0.64	0.64	0.63	0.62	0.61	0.61	0.60
0.6	0.1	0.4	0.64	0.63	0.63	0.62	0.61	0.60	0.60	0.59	0.58	0.58
0.6	0.1	0.5	0.64	0.63	0.62	0.61	0.61	0.60	0.59	0.59	0.58	0.58
0.6	0.2	0.1	0.88	0.87	0.87	0.86	0.86	0.86	0.85	0.85	0.84	0.84
0.6	0.2	0.2	0.81	0.81	0.80	0.80	0.79	0.79	0.78	0.78	0.77	0.77
0.6	0.2	0.3	0.78	0.78	0.77	0.77	0.76	0.75	0.75	0.75	0.74	0.74
0.6	0.2	0.4	0.77	0.77	0.76	0.76	0.75	0.75	0.74	0.74	0.73	0.73
0.6	0.2	0.5	0.78	0.77	0.77	0.76	0.76	0.75	0.75	0.75	0.74	0.74
0.6	0.3	0.1	0.94	0.94	0.93	0.93	0.93	0.93	0.92	0.92	0.92	0.92
0.6	0.3	0.2	0.91	0.90	0.90	0.90	0.89	0.89	0.89	0.89	0.88	0.88
0.6	0.3	0.3	0.89	0.89	0.89	0.88	0.88	0.88	0.87	0.87	0.87	0.87
0.6	0.3	0.4	0.89	0.89	0.88	0.88	0.88	0.88	0.87	0.87	0.87	0.87
0.6	0.3	0.5	0.89	0.89	0.89	0.89	0.88	0.88	0.88	0.88	0.88	0.88
0.6	0.4	0.1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.6	0.4	0.2	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.6	0.4	0.3	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.6	0.4	0.4	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.6	0.4	0.5	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.6	0.5	0.1	1.06	1.07	1.07	1.07	1.07	1.08	1.08	1.08	1.08	1.09
0.6	0.5	0.2	1.10	1.10	1.10	1.11	1.11	1.11	1.12	1.12	1.12	1.13
0.6	0.5	0.3	1.11	1.12	1.12	1.12	1.13	1.13	1.13	1.14	1.14	1.14
0.6	0.5	0.4	1.11	1.12	1.12	1.12	1.13	1.13	1.13	1.13	1.14	1.14
0.6	0.5	0.5	1.11	1.11	1.11	1.11	1.12	1.12	1.12	1.12	1.12	1.13
0.6	0.6	0.1	1.13	1.14	1.14	1.15	1.15	1.16	1.16	1.17	1.17	1.18
0.6	0.6	0.2	1.20	1.21	1.22	1.22	1.23	1.24	1.25	1.25	1.26	1.27
0.6	0.6	0.3	1.23	1.24	1.25	1.26	1.26	1.27	1.28	1.29	1.29	1.30
0.6	0.6	0.4	1.24	1.25	1.25	1.26	1.27	1.27	1.28	1.28	1.29	1.29
0.6	0.6	0.5	1.22	1.23	1.23	1.24	1.24	1.25	1.25	1.26	1.26	1.27
0.6	0.7	0.1	1.20	1.21	1.22	1.23	1.23	1.24	1.25	1.26	1.27	1.28
0.6	0.7	0.2	1.32	1.33	1.34	1.35	1.37	1.38	1.39	1.40	1.41	1.42
0.6	0.7	0.3	1.37	1.39	1.40	1.41	1.43	1.44	1.45	1.46	1.47	1.48
0.6	0.7	0.4	1.38	1.40	1.41	1.42	1.43	1.44	1.45	1.46	1.48	1.48
0.6	0.7	0.5	1.36	1.37	1.38	1.39	1.40	1.41	1.42	1.43	1.44	1.45
0.6	0.8	0.1	1.27	1.29	1.30	1.31	1.32	1.34	1.35	1.36	1.37	1.38
0.6	0.8	0.2	1.44	1.46	1.48	1.50	1.52	1.54	1.56	1.57	1.59	1.61
0.6	0.8	0.3	1.54	1.56	1.58	1.60	1.62	1.64	1.66	1.68	1.70	1.72
0.6	0.8	0.4	1.57	1.59	1.61	1.63	1.65	1.67	1.69	1.71	1.72	1.74
0.6	0.8	0.5	1.56	1.57	1.59	1.61	1.63	1.64	1.66	1.67	1.69	1.70
0.6	0.9	0.1	1.35	1.37	1.38	1.40	1.42	1.43	1.45	1.47	1.48	1.50
0.6	0.9	0.2	1.59	1.62	1.65	1.67	1.70	1.73	1.75	1.78	1.81	1.83
0.6	0.9	0.3	1.75	1.78	1.81	1.85	1.88	1.91	1.94	1.97	2.00	2.03
0.6	0.9	0.4	1.83	1.87	1.90	1.93	1.97	2.00	2.03	2.06	2.09	2.12
0.6	0.9	0.5	1.85	1.89	1.92	1.95	1.98	2.01	2.04	2.07	2.10	2.13
0.6	0.95	0.1	1.39	1.41	1.43	1.45	1.47	1.49	1.51	1.53	1.54	1.56
0.6	0.95	0.2	1.68	1.71	1.74	1.77	1.81	1.84	1.87	1.90	1.93	1.96
0.6	0.95	0.3	1.88	1.92	1.96	2.00	2.04	2.08	2.12	2.16	2.20	2.24
0.6	0.95	0.4	2.01	2.06	2.10	2.15	2.19	2.24	2.28	2.32	2.37	2.41
0.6	0.95	0.5	2.09	2.13	2.18	2.22	2.27	2.31	2.36	2.40	2.44	2.49
0.6	0.99	0.1	1.43	1.45	1.47	1.49	1.51	1.53	1.55	1.57	1.60	1.62
0.6	0.99	0.2	1.75	1.79	1.82	1.86	1.90	1.94	1.97	2.01	2.05	2.08
0.6	0.99	0.3	2.00	2.05	2.10	2.15	2.20	2.25	2.30	2.34	2.39	2.44
0.6	0.99	0.4	2.20	2.26	2.31	2.37	2.43	2.49	2.55	2.60	2.66	2.72
0.6	0.99	0.5	2.35	2.41	2.48	2.54	2.61	2.67	2.74	2.80	2.86	2.93

Table 3-6. Effect of non-differential misclassification of categorical data continued

P_N=the proportion exposed in the non-diseased group OR_X=Observed OR

		OR_T	1.05	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9
Specificity	**Sensitivity**	P_N	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X
0.7	0.1	0.1	1.00	0.99	0.98	0.97	0.97	0.96	0.95	0.94	0.94	0.93
0.7	0.1	0.2	0.99	0.98	0.97	0.95	0.94	0.93	0.91	0.90	0.89	0.88
0.7	0.1	0.3	0.99	0.98	0.96	0.94	0.92	0.90	0.89	0.87	0.86	0.84
0.7	0.1	0.4	0.99	0.97	0.95	0.93	0.91	0.89	0.87	0.85	0.84	0.82
0.7	0.1	0.5	0.98	0.97	0.94	0.92	0.90	0.88	0.86	0.84	0.83	0.81
0.7	0.2	0.1	1.00	1.00	0.99	0.99	0.98	0.98	0.98	0.97	0.97	0.96
0.7	0.2	0.2	1.00	0.99	0.98	0.98	0.97	0.96	0.96	0.95	0.95	0.94
0.7	0.2	0.3	0.99	0.99	0.98	0.97	0.96	0.95	0.95	0.94	0.93	0.93
0.7	0.2	0.4	0.99	0.99	0.98	0.97	0.96	0.95	0.94	0.93	0.93	0.92
0.7	0.2	0.5	0.99	0.99	0.98	0.97	0.96	0.95	0.94	0.93	0.93	0.92
0.7	0.3	0.1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.7	0.3	0.2	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.7	0.3	0.3	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.7	0.3	0.4	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.7	0.3	0.5	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.7	0.4	0.1	1.00	1.00	1.01	1.01	1.02	1.02	1.02	1.03	1.03	1.04
0.7	0.4	0.2	1.00	1.01	1.01	1.02	1.03	1.03	1.04	1.05	1.05	1.06
0.7	0.4	0.3	1.00	1.01	1.02	1.03	1.03	1.04	1.05	1.06	1.06	1.07
0.7	0.4	0.4	1.01	1.01	1.02	1.03	1.04	1.05	1.05	1.06	1.07	1.07
0.7	0.4	0.5	1.01	1.01	1.02	1.03	1.04	1.04	1.05	1.06	1.06	1.07
0.7	0.5	0.1	1.00	1.01	1.02	1.02	1.03	1.04	1.05	1.06	1.06	1.07
0.7	0.5	0.2	1.01	1.01	1.03	1.04	1.05	1.07	1.08	1.09	1.10	1.11
0.7	0.5	0.3	1.01	1.02	1.03	1.05	1.07	1.08	1.10	1.11	1.12	1.14
0.7	0.5	0.4	1.01	1.02	1.04	1.06	1.07	1.09	1.10	1.12	1.13	1.14
0.7	0.5	0.5	1.01	1.02	1.04	1.06	1.07	1.09	1.10	1.11	1.13	1.14
0.7	0.6	0.1	1.01	1.01	1.02	1.04	1.05	1.06	1.07	1.08	1.09	1.10
0.7	0.6	0.2	1.01	1.02	1.04	1.06	1.08	1.10	1.12	1.13	1.15	1.17
0.7	0.6	0.3	1.01	1.03	1.05	1.08	1.10	1.12	1.14	1.16	1.18	1.20
0.7	0.6	0.4	1.01	1.03	1.06	1.08	1.11	1.13	1.15	1.17	1.19	1.21
0.7	0.6	0.5	1.01	1.03	1.06	1.08	1.11	1.13	1.15	1.17	1.19	1.21
0.7	0.7	0.1	1.01	1.02	1.03	1.05	1.06	1.08	1.09	1.11	1.12	1.14
0.7	0.7	0.2	1.01	1.03	1.05	1.08	1.10	1.13	1.15	1.18	1.20	1.22
0.7	0.7	0.3	1.02	1.03	1.07	1.10	1.13	1.16	1.19	1.22	1.25	1.27
0.7	0.7	0.4	1.02	1.04	1.07	1.11	1.14	1.17	1.20	1.23	1.26	1.29
0.7	0.7	0.5	1.02	1.04	1.08	1.11	1.14	1.17	1.20	1.23	1.26	1.28
0.7	0.8	0.1	1.01	1.02	1.04	1.06	1.08	1.10	1.12	1.14	1.15	1.17
0.7	0.8	0.2	1.02	1.03	1.07	1.10	1.13	1.16	1.19	1.22	1.25	1.28
0.7	0.8	0.3	1.02	1.04	1.08	1.12	1.16	1.20	1.24	1.28	1.31	1.35
0.7	0.8	0.4	1.02	1.05	1.09	1.14	1.18	1.22	1.26	1.30	1.34	1.38
0.7	0.8	0.5	1.02	1.05	1.10	1.14	1.19	1.23	1.27	1.31	1.34	1.38
0.7	0.9	0.1	1.01	1.02	1.05	1.07	1.09	1.12	1.14	1.16	1.19	1.21
0.7	0.9	0.2	1.02	1.04	1.08	1.12	1.16	1.19	1.23	1.27	1.31	1.34
0.7	0.9	0.3	1.03	1.05	1.10	1.15	1.20	1.25	1.29	1.34	1.39	1.43
0.7	0.9	0.4	1.03	1.06	1.11	1.17	1.22	1.28	1.33	1.38	1.43	1.48
0.7	0.9	0.5	1.03	1.06	1.12	1.18	1.24	1.29	1.35	1.40	1.45	1.51
0.7	0.95	0.1	1.01	1.03	1.05	1.08	1.10	1.13	1.15	1.18	1.20	1.23
0.7	0.95	0.2	1.02	1.04	1.08	1.13	1.17	1.21	1.25	1.29	1.33	1.38
0.7	0.95	0.3	1.03	1.05	1.11	1.16	1.22	1.27	1.32	1.37	1.43	1.48
0.7	0.95	0.4	1.03	1.06	1.13	1.19	1.25	1.31	1.37	1.43	1.49	1.55
0.7	0.95	0.5	1.03	1.07	1.14	1.20	1.27	1.34	1.40	1.46	1.53	1.59
0.7	0.99	0.1	1.01	1.03	1.05	1.08	1.11	1.13	1.16	1.19	1.21	1.24
0.7	0.99	0.2	1.02	1.04	1.09	1.13	1.18	1.22	1.27	1.31	1.36	1.40
0.7	0.99	0.3	1.03	1.06	1.12	1.17	1.23	1.29	1.35	1.40	1.46	1.52
0.7	0.99	0.4	1.03	1.07	1.14	1.20	1.27	1.34	1.40	1.47	1.54	1.61
0.7	0.99	0.5	1.04	1.08	1.15	1.23	1.30	1.37	1.45	1.52	1.60	1.67

Table 3-6. Effect of non-differential misclassification of categorical data continued
P_N=the proportion exposed in the non-diseased group

		OR_T	2.0	2.1	2.2	2.3	2.4	2.5	2.6	2.7	2.8	2.9
Specificity	Sensitivity	P_N	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X
0.7	0.1	0.1	0.92	0.91	0.91	0.90	0.89	0.89	0.88	0.87	0.87	0.86
0.7	0.1	0.2	0.87	0.86	0.85	0.84	0.83	0.82	0.81	0.80	0.79	0.78
0.7	0.1	0.3	0.83	0.82	0.81	0.80	0.78	0.77	0.77	0.76	0.75	0.74
0.7	0.1	0.4	0.81	0.80	0.78	0.77	0.76	0.75	0.74	0.73	0.72	0.72
0.7	0.1	0.5	0.80	0.79	0.78	0.77	0.76	0.75	0.74	0.73	0.72	0.71
0.7	0.2	0.1	0.96	0.96	0.95	0.95	0.95	0.94	0.94	0.94	0.93	0.93
0.7	0.2	0.2	0.94	0.93	0.92	0.92	0.92	0.91	0.91	0.90	0.90	0.89
0.7	0.2	0.3	0.92	0.91	0.91	0.90	0.90	0.89	0.89	0.88	0.88	0.88
0.7	0.2	0.4	0.91	0.91	0.90	0.90	0.89	0.89	0.88	0.88	0.87	0.87
0.7	0.2	0.5	0.91	0.91	0.90	0.90	0.89	0.89	0.88	0.88	0.88	0.87
0.7	0.3	0.1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.7	0.3	0.2	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.7	0.3	0.3	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.7	0.3	0.4	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.7	0.3	0.5	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.7	0.4	0.1	1.04	1.04	1.05	1.05	1.05	1.06	1.06	1.06	1.07	1.07
0.7	0.4	0.2	1.06	1.07	1.07	1.08	1.08	1.09	1.09	1.10	1.10	1.10
0.7	0.4	0.3	1.07	1.08	1.09	1.09	1.10	1.10	1.11	1.11	1.12	1.12
0.7	0.4	0.4	1.08	1.08	1.09	1.09	1.10	1.10	1.11	1.11	1.12	1.12
0.7	0.4	0.5	1.08	1.08	1.08	1.09	1.09	1.10	1.10	1.10	1.11	1.11
0.7	0.5	0.1	1.08	1.08	1.09	1.10	1.11	1.11	1.12	1.13	1.13	1.14
0.7	0.5	0.2	1.12	1.13	1.14	1.15	1.16	1.17	1.18	1.19	1.20	1.21
0.7	0.5	0.3	1.15	1.16	1.17	1.18	1.19	1.20	1.21	1.22	1.23	1.24
0.7	0.5	0.4	1.15	1.17	1.18	1.19	1.20	1.21	1.22	1.22	1.23	1.24
0.7	0.5	0.5	1.15	1.16	1.17	1.18	1.18	1.19	1.20	1.21	1.21	1.22
0.7	0.6	0.1	1.12	1.13	1.14	1.15	1.16	1.17	1.18	1.19	1.20	1.21
0.7	0.6	0.2	1.19	1.20	1.22	1.23	1.25	1.26	1.28	1.29	1.31	1.32
0.7	0.6	0.3	1.22	1.24	1.26	1.27	1.29	1.31	1.32	1.34	1.35	1.37
0.7	0.6	0.4	1.23	1.25	1.27	1.28	1.30	1.31	1.33	1.34	1.36	1.37
0.7	0.6	0.5	1.22	1.24	1.25	1.27	1.28	1.29	1.31	1.32	1.33	1.34
0.7	0.7	0.1	1.15	1.17	1.18	1.20	1.21	1.23	1.24	1.25	1.27	1.28
0.7	0.7	0.2	1.25	1.27	1.29	1.31	1.33	1.36	1.38	1.40	1.42	1.44
0.7	0.7	0.3	1.30	1.32	1.35	1.37	1.40	1.42	1.44	1.46	1.49	1.51
0.7	0.7	0.4	1.32	1.34	1.37	1.39	1.41	1.43	1.46	1.48	1.50	1.52
0.7	0.7	0.5	1.31	1.33	1.35	1.37	1.39	1.41	1.43	1.45	1.47	1.48
0.7	0.8	0.1	1.19	1.21	1.23	1.25	1.27	1.28	1.30	1.32	1.34	1.36
0.7	0.8	0.2	1.31	1.34	1.37	1.40	1.43	1.45	1.48	1.51	1.54	1.56
0.7	0.8	0.3	1.38	1.42	1.45	1.48	1.52	1.55	1.58	1.61	1.64	1.67
0.7	0.8	0.4	1.41	1.45	1.48	1.52	1.55	1.58	1.61	1.64	1.67	1.70
0.7	0.8	0.5	1.41	1.45	1.48	1.51	1.54	1.57	1.60	1.62	1.65	1.67
0.7	0.9	0.1	1.23	1.25	1.28	1.30	1.32	1.34	1.37	1.39	1.41	1.43
0.7	0.9	0.2	1.38	1.42	1.45	1.49	1.53	1.56	1.60	1.63	1.67	1.70
0.7	0.9	0.3	1.48	1.52	1.57	1.61	1.65	1.70	1.74	1.78	1.82	1.86
0.7	0.9	0.4	1.53	1.58	1.63	1.68	1.72	1.77	1.81	1.86	1.90	1.95
0.7	0.9	0.5	1.56	1.60	1.65	1.70	1.74	1.79	1.83	1.88	1.92	1.96
0.7	0.95	0.1	1.25	1.28	1.30	1.32	1.35	1.37	1.40	1.42	1.45	1.47
0.7	0.95	0.2	1.42	1.46	1.50	1.54	1.58	1.62	1.66	1.70	1.74	1.78
0.7	0.95	0.3	1.53	1.58	1.63	1.68	1.73	1.78	1.83	1.88	1.93	1.98
0.7	0.95	0.4	1.61	1.66	1.72	1.78	1.83	1.89	1.94	2.00	2.05	2.11
0.7	0.95	0.5	1.65	1.71	1.77	1.83	1.89	1.95	2.00	2.06	2.11	2.17
0.7	0.99	0.1	1.27	1.29	1.32	1.35	1.37	1.40	1.43	1.45	1.48	1.51
0.7	0.99	0.2	1.45	1.49	1.54	1.58	1.62	1.67	1.71	1.76	1.80	1.85
0.7	0.99	0.3	1.58	1.63	1.69	1.75	1.80	1.86	1.92	1.98	2.03	2.09
0.7	0.99	0.4	1.67	1.74	1.80	1.87	1.94	2.00	2.07	2.13	2.20	2.27
0.7	0.99	0.5	1.74	1.82	1.89	1.96	2.03	2.11	2.18	2.25	2.32	2.39

Table 3-6. Effect of non-differential misclassification of categorical data continued
P_N=the proportion exposed in the non-diseased group

		OR_T	3.0	3.1	3.2	3.3	3.4	3.5	3.6	3.7	3.8	3.9
Specificity	**Sensitivity**	P_N	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X
0.7	0.1	0.1	0.86	0.85	0.85	0.84	0.84	0.83	0.82	0.82	0.81	0.81
0.7	0.1	0.2	0.78	0.77	0.76	0.75	0.75	0.74	0.74	0.73	0.72	0.72
0.7	0.1	0.3	0.73	0.72	0.72	0.71	0.70	0.70	0.69	0.68	0.68	0.67
0.7	0.1	0.4	0.71	0.70	0.69	0.69	0.68	0.68	0.67	0.66	0.66	0.65
0.7	0.1	0.5	0.71	0.70	0.69	0.69	0.68	0.68	0.67	0.67	0.66	0.66
0.7	0.2	0.1	0.93	0.93	0.92	0.92	0.92	0.91	0.91	0.91	0.91	0.90
0.7	0.2	0.2	0.89	0.89	0.88	0.88	0.88	0.87	0.87	0.87	0.86	0.86
0.7	0.2	0.3	0.87	0.87	0.86	0.86	0.86	0.85	0.85	0.85	0.84	0.84
0.7	0.2	0.4	0.87	0.86	0.86	0.86	0.85	0.85	0.85	0.84	0.84	0.84
0.7	0.2	0.5	0.87	0.87	0.87	0.86	0.86	0.86	0.85	0.85	0.85	0.85
0.7	0.3	0.1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.7	0.3	0.2	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.7	0.3	0.3	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.7	0.3	0.4	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.7	0.3	0.5	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.7	0.4	0.1	1.07	1.07	1.08	1.08	1.08	1.09	1.09	1.09	1.09	1.10
0.7	0.4	0.2	1.11	1.11	1.12	1.12	1.12	1.13	1.13	1.13	1.14	1.14
0.7	0.4	0.3	1.12	1.13	1.13	1.14	1.14	1.14	1.15	1.15	1.15	1.15
0.7	0.4	0.4	1.12	1.13	1.13	1.13	1.14	1.14	1.14	1.15	1.15	1.15
0.7	0.4	0.5	1.11	1.12	1.12	1.12	1.13	1.13	1.13	1.13	1.13	1.14
0.7	0.5	0.1	1.14	1.15	1.16	1.16	1.17	1.17	1.18	1.19	1.19	1.20
0.7	0.5	0.2	1.22	1.23	1.24	1.24	1.25	1.26	1.27	1.27	1.28	1.29
0.7	0.5	0.3	1.25	1.26	1.26	1.27	1.28	1.29	1.29	1.30	1.31	1.31
0.7	0.5	0.4	1.25	1.26	1.26	1.27	1.28	1.28	1.29	1.29	1.30	1.31
0.7	0.5	0.5	1.23	1.23	1.24	1.24	1.25	1.26	1.26	1.26	1.27	1.27
0.7	0.6	0.1	1.22	1.23	1.24	1.25	1.26	1.27	1.27	1.28	1.29	1.30
0.7	0.6	0.2	1.33	1.35	1.36	1.37	1.38	1.40	1.41	1.42	1.43	1.44
0.7	0.6	0.3	1.38	1.39	1.41	1.42	1.43	1.44	1.46	1.47	1.48	1.49
0.7	0.6	0.4	1.38	1.39	1.40	1.42	1.43	1.44	1.45	1.46	1.47	1.48
0.7	0.6	0.5	1.35	1.36	1.37	1.38	1.39	1.40	1.40	1.41	1.42	1.43
0.7	0.7	0.1	1.29	1.31	1.32	1.33	1.35	1.36	1.37	1.39	1.40	1.41
0.7	0.7	0.2	1.46	1.47	1.49	1.51	1.53	1.55	1.66	1.58	1.60	1.62
0.7	0.7	0.3	1.53	1.55	1.57	1.58	1.60	1.62	1.64	1.66	1.67	1.69
0.7	0.7	0.4	1.54	1.55	1.57	1.59	1.60	1.62	1.64	1.65	1.67	1.68
0.7	0.7	0.5	1.50	1.52	1.53	1.54	1.56	1.57	1.58	1.60	1.61	1.62
0.7	0.8	0.1	1.37	1.39	1.41	1.42	1.44	1.46	1.48	1.49	1.51	1.53
0.7	0.8	0.2	1.59	1.61	1.64	1.66	1.69	1.71	1.74	1.76	1.79	1.81
0.7	0.8	0.3	1.70	1.72	1.75	1.78	1.81	1.83	1.86	1.88	1.91	1.93
0.7	0.8	0.4	1.73	1.75	1.78	1.81	1.83	1.86	1.88	1.91	1.93	1.95
0.7	0.8	0.5	1.70	1.72	1.75	1.77	1.79	1.81	1.83	1.85	1.87	1.89
0.7	0.9	0.1	1.45	1.48	1.50	1.52	1.54	1.56	1.59	1.61	1.63	1.65
0.7	0.9	0.2	1.74	1.77	1.81	1.84	1.87	1.91	1.94	1.97	2.01	2.04
0.7	0.9	0.3	1.91	1.95	1.99	2.03	2.06	2.10	2.14	2.18	2.22	2.25
0.7	0.9	0.4	1.99	2.03	2.07	2.11	2.15	2.19	2.23	2.27	2.31	2.34
0.7	0.9	0.5	2.00	2.04	2.08	2.12	2.15	2.19	2.23	2.26	2.30	2.33
0.7	0.95	0.1	1.50	1.52	1.55	1.57	1.59	1.62	1.64	1.67	1.69	1.72
0.7	0.95	0.2	1.82	1.86	1.90	1.94	1.98	2.02	2.06	2.09	2.13	2.17
0.7	0.95	0.3	2.03	2.08	2.13	2.18	2.22	2.27	2.32	2.36	2.41	2.46
0.7	0.95	0.4	2.16	2.21	2.27	2.32	2.37	2.42	2.47	2.52	2.57	2.62
0.7	0.95	0.5	2.22	2.28	2.33	2.38	2.43	2.49	2.54	2.59	2.64	2.68
0.7	0.99	0.1	1.53	1.56	1.58	1.61	1.64	1.66	1.69	1.72	1.74	1.77
0.7	0.99	0.2	1.89	1.93	1.98	2.02	2.07	2.11	2.16	2.20	2.24	2.29
0.7	0.99	0.3	2.15	2.20	2.26	2.31	2.37	2.43	2.48	2.54	2.60	2.65
0.7	0.99	0.4	2.33	2.40	2.46	2.53	2.59	2.66	2.72	2.79	2.85	2.91
0.7	0.99	0.5	2.47	2.54	2.61	2.68	2.75	2.82	2.89	2.96	3.03	3.10

Table 3-6. Effect of non-differential misclassification of categorical data continued

P_N=the proportion exposed in the non-diseased group **OR_X=Observed OR**

		OR_T	1.05	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9
Specificity	**Sensitivity**	P_N	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X
0.8	0.1	0.1	1.00	0.99	0.99	0.98	0.98	0.97	0.97	0.96	0.96	0.95
0.8	0.1	0.2	0.99	0.99	0.98	0.97	0.96	0.95	0.94	0.93	0.93	0.92
0.8	0.1	0.3	0.99	0.99	0.97	0.96	0.95	0.94	0.93	0.92	0.91	0.90
0.8	0.1	0.4	0.99	0.98	0.97	0.95	0.94	0.93	0.91	0.90	0.89	0.88
0.8	0.1	0.5	0.99	0.98	0.96	0.95	0.94	0.92	0.91	0.90	0.89	0.88
0.8	0.2	0.1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.8	0.2	0.2	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.8	0.2	0.3	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.8	0.2	0.4	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.8	0.2	0.5	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.8	0.3	0.1	1.00	1.01	1.01	1.02	1.02	1.03	1.03	1.04	1.04	1.05
0.8	0.3	0.2	1.00	1.01	1.02	1.03	1.03	1.04	1.05	1.06	1.07	1.07
0.8	0.3	0.3	1.01	1.01	1.02	1.03	1.04	1.05	1.06	1.07	1.08	1.09
0.8	0.3	0.4	1.01	1.01	1.02	1.04	1.05	1.06	1.06	1.07	1.08	1.09
0.8	0.3	0.5	1.01	1.01	1.02	1.04	1.04	1.05	1.06	1.07	1.08	1.08
0.8	0.4	0.1	1.01	1.01	1.02	1.03	1.04	1.05	1.06	1.07	1.08	1.09
0.8	0.4	0.2	1.01	1.02	1.03	1.05	1.07	1.08	1.10	1.11	1.12	1.14
0.8	0.4	0.3	1.01	1.02	1.04	1.06	1.08	1.10	1.11	1.13	1.15	1.16
0.8	0.4	0.4	1.01	1.02	1.04	1.06	1.08	1.10	1.12	1.14	1.15	1.16
0.8	0.4	0.5	1.01	1.02	1.04	1.06	1.08	1.10	1.11	1.13	1.14	1.15
0.8	0.5	0.1	1.01	1.02	1.03	1.04	1.06	1.07	1.09	1.10	1.12	1.13
0.8	0.5	0.2	1.01	1.02	1.05	1.07	1.09	1.12	1.14	1.16	1.18	1.20
0.8	0.5	0.3	1.02	1.03	1.06	1.09	1.11	1.14	1.16	1.19	1.21	1.23
0.8	0.5	0.4	1.02	1.03	1.06	1.09	1.12	1.14	1.17	1.19	1.21	1.24
0.8	0.5	0.5	1.02	1.03	1.06	1.09	1.11	1.14	1.16	1.18	1.20	1.22
0.8	0.6	0.1	1.01	1.02	1.04	1.06	1.08	1.10	1.11	1.13	1.15	1.17
0.8	0.6	0.2	1.02	1.03	1.06	1.09	1.12	1.15	1.18	1.21	1.23	1.26
0.8	0.6	0.3	1.02	1.04	1.07	1.11	1.14	1.18	1.21	1.24	1.27	1.30
0.8	0.6	0.4	1.02	1.04	1.08	1.12	1.15	1.19	1.22	1.25	1.28	1.31
0.8	0.6	0.5	1.02	1.04	1.08	1.11	1.15	1.18	1.21	1.24	1.26	1.29
0.8	0.7	0.1	1.01	1.02	1.05	1.07	1.09	1.12	1.14	1.16	1.19	1.21
0.8	0.7	0.2	1.02	1.04	1.07	1.11	1.15	1.18	1.22	1.25	1.29	1.32
0.8	0.7	0.3	1.02	1.05	1.09	1.13	1.17	1.22	1.26	1.29	1.33	1.37
0.8	0.7	0.4	1.02	1.05	1.10	1.14	1.19	1.23	1.27	1.31	1.34	1.38
0.8	0.7	0.5	1.02	1.05	1.10	1.14	1.18	1.22	1.26	1.30	1.33	1.36
0.8	0.8	0.1	1.01	1.03	1.06	1.08	1.11	1.14	1.17	1.19	1.22	1.25
0.8	0.8	0.2	1.02	1.04	1.09	1.13	1.17	1.21	1.26	1.30	1.34	1.38
0.8	0.8	0.3	1.03	1.05	1.10	1.16	1.21	1.26	1.30	1.35	1.40	1.44
0.8	0.8	0.4	1.03	1.06	1.11	1.17	1.22	1.27	1.32	1.37	1.42	1.47
0.8	0.8	0.5	1.03	1.06	1.12	1.17	1.22	1.27	1.32	1.37	1.41	1.46
0.8	0.9	0.1	1.02	1.03	1.06	1.10	1.13	1.16	1.19	1.22	1.25	1.28
0.8	0.9	0.2	1.02	1.05	1.10	1.15	1.20	1.25	1.29	1.34	1.39	1.44
0.8	0.9	0.3	1.03	1.06	1.12	1.18	1.24	1.30	1.35	1.41	1.47	1.52
0.8	0.9	0.4	1.03	1.07	1.13	1.20	1.26	1.32	1.39	1.45	1.51	1.57
0.8	0.9	0.5	1.04	1.07	1.14	1.21	1.27	1.33	1.40	1.46	1.52	1.58
0.8	0.95	0.1	1.02	1.03	1.07	1.10	1.14	1.17	1.20	1.24	1.27	1.30
0.8	0.95	0.2	1.03	1.05	1.11	1.16	1.21	1.26	1.31	1.37	1.42	1.47
0.8	0.95	0.3	1.03	1.06	1.13	1.19	1.26	1.32	1.38	1.44	1.51	1.57
0.8	0.95	0.4	1.04	1.07	1.14	1.21	1.28	1.35	1.42	1.49	1.56	1.63
0.8	0.95	0.5	1.04	1.08	1.15	1.23	1.30	1.37	1.44	1.52	1.59	1.66
0.8	0.99	0.1	1.02	1.04	1.07	1.11	1.14	1.18	1.21	1.25	1.28	1.32
0.8	0.99	0.2	1.03	1.05	1.11	1.16	1.22	1.27	1.33	1.38	1.44	1.49
0.8	0.99	0.3	1.03	1.07	1.13	1.20	1.27	1.34	1.40	1.47	1.54	1.60
0.8	0.99	0.4	1.04	1.08	1.15	1.23	1.30	1.38	1.45	1.53	1.60	1.68
0.8	0.99	0.5	1.04	1.08	1.16	1.24	1.33	1.41	1.49	1.57	1.65	1.73

Table 3-6. Effect of non-differential misclassification of categorical data continued
P_N=the proportion exposed in the non-diseased group

		OR_T	2.0	2.1	2.2	2.3	2.4	2.5	2.6	2.7	2.8	2.9
Specificity	Sensitivity	P_N	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X
0.8	0.1	0.1	0.95	0.94	0.94	0.93	0.93	0.92	0.92	0.92	0.91	0.91
0.8	0.1	0.2	0.91	0.90	0.90	0.89	0.88	0.88	0.87	0.87	0.86	0.85
0.8	0.1	0.3	0.89	0.88	0.87	0.86	0.86	0.85	0.84	0.84	0.83	0.83
0.8	0.1	0.4	0.88	0.87	0.86	0.85	0.84	0.84	0.83	0.82	0.82	0.81
0.8	0.1	0.5	0.87	0.86	0.86	0.85	0.84	0.84	0.83	0.82	0.82	0.81
0.8	0.2	0.1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.8	0.2	0.2	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.8	0.2	0.3	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.8	0.2	0.4	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.8	0.2	0.5	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.8	0.3	0.1	1.05	1.05	1.06	1.06	1.07	1.07	1.08	1.08	1.08	1.09
0.8	0.3	0.2	1.08	1.09	1.09	1.10	1.10	1.11	1.12	1.12	1.13	1.13
0.8	0.3	0.3	1.09	1.10	1.11	1.11	1.12	1.13	1.13	1.14	1.14	1.15
0.8	0.3	0.4	1.10	1.10	1.11	1.12	1.12	1.13	1.13	1.14	1.14	1.15
0.8	0.3	0.5	1.09	1.10	1.10	1.11	1.11	1.12	1.12	1.13	1.13	1.13
0.8	0.4	0.1	1.10	1.11	1.12	1.12	1.13	1.14	1.15	1.16	1.17	1.17
0.8	0.4	0.2	1.15	1.16	1.18	1.19	1.20	1.21	1.22	1.24	1.25	1.26
0.8	0.4	0.3	1.18	1.19	1.20	1.22	1.23	1.24	1.25	1.26	1.27	1.28
0.8	0.4	0.4	1.18	1.19	1.20	1.22	1.23	1.24	1.25	1.26	1.27	1.28
0.8	0.4	0.5	1.17	1.18	1.19	1.20	1.21	1.22	1.23	1.23	1.24	1.25
0.8	0.5	0.1	1.14	1.16	1.17	1.18	1.20	1.21	1.22	1.23	1.25	1.26
0.8	0.5	0.2	1.22	1.24	1.26	1.28	1.29	1.31	1.33	1.34	1.36	1.38
0.8	0.5	0.3	1.25	1.27	1.29	1.31	1.33	1.35	1.37	1.38	1.40	1.41
0.8	0.5	0.4	1.26	1.28	1.29	1.31	1.33	1.34	1.36	1.38	1.39	1.40
0.8	0.5	0.5	1.24	1.25	1.27	1.29	1.30	1.31	1.33	1.34	1.35	1.36
0.8	0.6	0.1	1.19	1.21	1.22	1.24	1.26	1.27	1.29	1.31	1.32	1.34
0.8	0.6	0.2	1.29	1.31	1.34	1.36	1.38	1.41	1.43	1.45	1.48	1.50
0.8	0.6	0.3	1.33	1.36	1.38	1.41	1.43	1.46	1.48	1.50	1.53	1.55
0.8	0.6	0.4	1.33	1.36	1.38	1.41	1.43	1.45	1.48	1.50	1.52	1.54
0.8	0.6	0.5	1.31	1.34	1.36	1.38	1.40	1.42	1.43	1.45	1.47	1.48
0.8	0.7	0.1	1.23	1.25	1.27	1.30	1.32	1.34	1.36	1.38	1.40	1.42
0.8	0.7	0.2	1.35	1.38	1.41	1.45	1.48	1.51	1.54	1.57	1.59	1.62
0.8	0.7	0.3	1.41	1.44	1.47	1.51	1.54	1.57	1.60	1.64	1.67	1.69
0.8	0.7	0.4	1.42	1.45	1.48	1.52	1.55	1.58	1.61	1.63	1.66	1.69
0.8	0.7	0.5	1.40	1.43	1.46	1.48	1.51	1.54	1.56	1.59	1.61	1.63
0.8	0.8	0.1	1.27	1.30	1.33	1.35	1.38	1.40	1.43	1.46	1.48	1.51
0.8	0.8	0.2	1.42	1.46	1.49	1.53	1.57	1.61	1.65	1.68	1.72	1.75
0.8	0.8	0.3	1.49	1.53	1.57	1.62	1.66	1.70	1.74	1.78	1.82	1.86
0.8	0.8	0.4	1.51	1.56	1.60	1.64	1.68	1.72	1.76	1.80	1.84	1.87
0.8	0.8	0.5	1.50	1.54	1.58	1.62	1.66	1.69	1.73	1.76	1.79	1.83
0.8	0.9	0.1	1.32	1.35	1.38	1.41	1.44	1.47	1.50	1.53	1.56	1.59
0.8	0.9	0.2	1.48	1.53	1.58	1.62	1.67	1.72	1.76	1.81	1.85	1.90
0.8	0.9	0.3	1.58	1.63	1.69	1.74	1.79	1.85	1.90	1.95	2.00	2.05
0.8	0.9	0.4	1.63	1.68	1.74	1.80	1.85	1.91	1.96	2.01	2.06	2.12
0.8	0.9	0.5	1.64	1.69	1.75	1.80	1.86	1.91	1.96	2.01	2.06	2.11
0.8	0.95	0.1	1.34	1.37	1.40	1.44	1.47	1.50	1.54	1.57	1.60	1.63
0.8	0.95	0.2	1.52	1.57	1.62	1.67	1.72	1.77	1.82	1.87	1.92	1.97
0.8	0.95	0.3	1.63	1.69	1.75	1.81	1.87	1.93	1.99	2.05	2.11	2.17
0.8	0.95	0.4	1.69	1.76	1.82	1.89	1.95	2.02	2.08	2.15	2.21	2.27
0.8	0.95	0.5	1.72	1.79	1.86	1.93	1.99	2.06	2.12	2.19	2.25	2.31
0.8	0.99	0.1	1.35	1.39	1.42	1.46	1.49	1.53	1.56	1.60	1.63	1.67
0.8	0.99	0.2	1.55	1.60	1.66	1.71	1.77	1.82	1.88	1.93	1.98	2.04
0.8	0.99	0.3	1.67	1.74	1.80	1.87	1.94	2.00	2.07	2.14	2.20	2.27
0.8	0.99	0.4	1.75	1.83	1.90	1.98	2.05	2.12	2.20	2.27	2.35	2.42
0.8	0.99	0.5	1.81	1.89	1.97	2.05	2.13	2.21	2.29	2.36	2.44	2.52

Table 3-6. Effect of non-differential misclassification of categorical data continued
P_N=the proportion exposed in the non-diseased group

		OR_T	3.0	3.1	3.2	3.3	3.4	3.5	3.6	3.7	3.8	3.9
Specificity	**Sensitivity**	P_N	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X
0.8	0.1	0.1	0.90	0.90	0.90	0.89	0.89	0.70	0.88	0.88	0.88	0.87
0.8	0.1	0.2	0.85	0.84	0.84	0.83	0.83	0.58	0.82	0.82	0.81	0.81
0.8	0.1	0.3	0.82	0.81	0.81	0.80	0.80	0.52	0.79	0.79	0.78	0.78
0.8	0.1	0.4	0.81	0.80	0.80	0.79	0.79	0.50	0.78	0.78	0.77	0.77
0.8	0.1	0.5	0.81	0.81	0.80	0.80	0.79	0.50	0.79	0.78	0.78	0.78
0.8	0.2	0.1	1.00	1.00	1.00	1.00	1.00	0.75	1.00	1.00	1.00	1.00
0.8	0.2	0.2	1.00	1.00	1.00	1.00	1.00	0.65	1.00	1.00	1.00	1.00
0.8	0.2	0.3	1.00	1.00	1.00	1.00	1.00	0.61	1.00	1.00	1.00	1.00
0.8	0.2	0.4	1.00	1.00	1.00	1.00	1.00	0.60	1.00	1.00	1.00	1.00
0.8	0.2	0.5	1.00	1.00	1.00	1.00	1.00	0.61	1.00	1.00	1.00	1.00
0.8	0.3	0.1	1.09	1.10	1.10	1.10	1.11	0.80	1.11	1.12	1.12	1.13
0.8	0.3	0.2	1.14	1.14	1.15	1.15	1.16	0.73	1.17	1.17	1.17	1.18
0.8	0.3	0.3	1.15	1.16	1.16	1.17	1.17	0.70	1.18	1.18	1.19	1.19
0.8	0.3	0.4	1.15	1.16	1.16	1.16	1.17	0.69	1.17	1.18	1.18	1.18
0.8	0.3	0.5	1.14	1.14	1.14	1.15	1.15	0.71	1.16	1.16	1.16	1.16
0.8	0.4	0.1	1.18	1.19	1.20	1.20	1.21	0.86	1.23	1.23	1.24	1.25
0.8	0.4	0.2	1.27	1.28	1.29	1.30	1.31	0.81	1.32	1.33	1.34	1.35
0.8	0.4	0.3	1.29	1.30	1.31	1.32	1.33	0.79	1.35	1.36	1.36	1.37
0.8	0.4	0.4	1.29	1.29	1.30	1.31	1.32	0.79	1.33	1.34	1.34	1.35
0.8	0.4	0.5	1.26	1.26	1.27	1.28	1.28	0.80	1.29	1.30	1.30	1.31
0.8	0.5	0.1	1.27	1.28	1.29	1.31	1.32	0.93	1.34	1.35	1.36	1.37
0.8	0.5	0.2	1.39	1.41	1.42	1.44	1.45	0.90	1.48	1.49	1.51	1.52
0.8	0.5	0.3	1.43	1.45	1.46	1.47	1.49	0.89	1.51	1.53	1.54	1.55
0.8	0.5	0.4	1.42	1.43	1.44	1.45	1.47	0.89	1.49	1.50	1.51	1.52
0.8	0.5	0.5	1.37	1.38	1.39	1.40	1.41	0.89	1.43	1.44	1.44	1.45
0.8	0.6	0.1	1.36	1.37	1.39	1.40	1.42	1.00	1.45	1.47	1.48	1.50
0.8	0.6	0.2	1.52	1.54	1.56	1.58	1.60	1.00	1.64	1.66	1.68	1.70
0.8	0.6	0.3	1.57	1.59	1.61	1.63	1.65	1.00	1.69	1.71	1.72	1.74
0.8	0.6	0.4	1.56	1.57	1.59	1.61	1.63	1.00	1.66	1.67	1.69	1.70
0.8	0.6	0.5	1.50	1.51	1.53	1.54	1.56	1.00	1.58	1.59	1.60	1.61
0.8	0.7	0.1	1.44	1.46	1.49	1.51	1.53	1.08	1.57	1.58	1.60	1.62
0.8	0.7	0.2	1.65	1.68	1.71	1.73	1.76	1.12	1.81	1.84	1.86	1.88
0.8	0.7	0.3	1.72	1.75	1.78	1.81	1.83	1.14	1.88	1.91	1.93	1.95
0.8	0.7	0.4	1.71	1.74	1.76	1.79	1.81	1.14	1.86	1.88	1.90	1.92
0.8	0.7	0.5	1.65	1.67	1.69	1.71	1.73	1.13	1.77	1.79	1.80	1.82
0.8	0.8	0.1	1.53	1.56	1.58	1.61	1.63	1.17	1.68	1.71	1.73	1.75
0.8	0.8	0.2	1.79	1.82	1.86	1.89	1.93	1.27	1.99	2.03	2.06	2.09
0.8	0.8	0.3	1.90	1.93	1.97	2.01	2.04	1.32	2.11	2.14	2.18	2.21
0.8	0.8	0.4	1.91	1.94	1.98	2.01	2.04	1.34	2.11	2.14	2.17	2.20
0.8	0.8	0.5	1.86	1.89	1.92	1.95	1.97	1.32	2.03	2.05	2.08	2.10
0.8	0.9	0.1	1.62	1.65	1.68	1.71	1.74	1.27	1.80	1.83	1.86	1.89
0.8	0.9	0.2	1.94	1.99	2.03	2.07	2.12	1.47	2.20	2.25	2.29	2.33
0.8	0.9	0.3	2.10	2.15	2.20	2.25	2.30	1.59	2.40	2.44	2.49	2.54
0.8	0.9	0.4	2.17	2.22	2.27	2.32	2.36	1.66	2.46	2.50	2.55	2.60
0.8	0.9	0.5	2.16	2.20	2.25	2.30	2.34	1.67	2.43	2.47	2.51	2.55
0.8	0.95	0.1	1.67	1.70	1.73	1.77	1.80	1.33	1.86	1.90	1.93	1.96
0.8	0.95	0.2	2.02	2.07	2.12	2.17	2.22	1.59	2.32	2.37	2.42	2.46
0.8	0.95	0.3	2.23	2.28	2.34	2.40	2.46	1.78	2.57	2.63	2.68	2.74
0.8	0.95	0.4	2.33	2.39	2.46	2.52	2.58	1.92	2.70	2.75	2.81	2.87
0.8	0.95	0.5	2.37	2.43	2.49	2.55	2.61	1.98	2.73	2.79	2.84	2.90
0.8	0.99	0.1	1.70	1.74	1.78	1.81	1.85	1.38	1.92	1.95	1.99	2.02
0.8	0.99	0.2	2.09	2.15	2.20	2.26	2.31	1.70	2.42	2.47	2.53	2.58
0.8	0.99	0.3	2.33	2.40	2.47	2.53	2.60	1.98	2.73	2.79	2.86	2.93
0.8	0.99	0.4	2.49	2.57	2.64	2.71	2.79	2.22	2.93	3.01	3.08	3.15
0.8	0.99	0.5	2.60	2.68	2.76	2.83	2.91	2.41	3.06	3.14	3.22	3.29

Table 3-6. Effect of non-differential misclassification of categorical data continued

P_N=the proportion exposed in the non-diseased group $\quad$ OR_X=Observed OR

		OR_T	1.05	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9
Specificity	**Sensitivity**	P_N	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X
0.9	0.1	0.1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.9	0.1	0.2	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.9	0.1	0.3	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.9	0.1	0.4	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.9	0.1	0.5	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.9	0.2	0.1	1.00	1.01	1.02	1.03	1.04	1.04	1.05	1.06	1.07	1.08
0.9	0.2	0.2	1.01	1.01	1.03	1.04	1.06	1.07	1.08	1.09	1.11	1.12
0.9	0.2	0.3	1.01	1.02	1.04	1.05	1.07	1.08	1.10	1.11	1.12	1.13
0.9	0.2	0.4	1.01	1.02	1.04	1.05	1.07	1.08	1.10	1.11	1.12	1.13
0.9	0.2	0.5	1.01	1.02	1.04	1.05	1.07	1.08	1.09	1.10	1.11	1.12
0.9	0.3	0.1	1.01	1.02	1.03	1.05	1.07	1.08	1.10	1.11	1.13	1.14
0.9	0.3	0.2	1.01	1.03	1.05	1.08	1.10	1.12	1.15	1.17	1.19	1.21
0.9	0.3	0.3	1.02	1.03	1.06	1.09	1.11	1.14	1.16	1.19	1.21	1.23
0.9	0.3	0.4	1.02	1.03	1.06	1.09	1.11	1.14	1.16	1.18	1.20	1.22
0.9	0.3	0.5	1.02	1.03	1.06	1.08	1.11	1.13	1.15	1.17	1.19	1.20
0.9	0.4	0.1	1.01	1.02	1.05	1.07	1.09	1.12	1.14	1.16	1.18	1.20
0.9	0.4	0.2	1.02	1.04	1.07	1.10	1.14	1.17	1.20	1.23	1.26	1.28
0.9	0.4	0.3	1.02	1.04	1.08	1.12	1.15	1.18	1.22	1.25	1.28	1.31
0.9	0.4	0.4	1.02	1.04	1.08	1.12	1.15	1.18	1.21	1.24	1.27	1.30
0.9	0.4	0.5	1.02	1.04	1.07	1.11	1.14	1.17	1.19	1.22	1.24	1.26
0.9	0.5	0.1	1.01	1.03	1.06	1.09	1.12	1.15	1.17	1.20	1.23	1.26
0.9	0.5	0.2	1.02	1.04	1.08	1.13	1.17	1.20	1.24	1.28	1.32	1.35
0.9	0.5	0.3	1.02	1.05	1.09	1.14	1.18	1.22	1.26	1.30	1.34	1.38
0.9	0.5	0.4	1.02	1.05	1.09	1.14	1.18	1.22	1.26	1.29	1.33	1.36
0.9	0.5	0.5	1.02	1.05	1.09	1.13	1.17	1.20	1.24	1.27	1.30	1.32
0.9	0.6	0.1	1.02	1.04	1.07	1.10	1.14	1.17	1.21	1.24	1.27	1.30
0.9	0.6	0.2	1.02	1.05	1.10	1.15	1.19	1.24	1.28	1.33	1.37	1.41
0.9	0.6	0.3	1.03	1.06	1.11	1.16	1.21	1.26	1.31	1.35	1.40	1.44
0.9	0.6	0.4	1.03	1.06	1.11	1.16	1.21	1.26	1.30	1.34	1.39	1.43
0.9	0.6	0.5	1.03	1.05	1.10	1.15	1.20	1.24	1.28	1.32	1.35	1.39
0.9	0.7	0.1	1.02	1.04	1.08	1.12	1.16	1.20	1.24	1.27	1.31	1.35
0.9	0.7	0.2	1.03	1.06	1.11	1.16	1.22	1.27	1.32	1.37	1.42	1.47
0.9	0.7	0.3	1.03	1.06	1.12	1.18	1.24	1.29	1.35	1.40	1.45	1.51
0.9	0.7	0.4	1.03	1.06	1.12	1.18	1.24	1.29	1.35	1.40	1.45	1.50
0.9	0.7	0.5	1.03	1.06	1.12	1.17	1.23	1.28	1.33	1.37	1.42	1.46
0.9	0.8	0.1	1.02	1.04	1.09	1.13	1.18	1.22	1.26	1.31	1.35	1.39
0.9	0.8	0.2	1.03	1.06	1.12	1.18	1.24	1.30	1.36	1.41	1.47	1.53
0.9	0.8	0.3	1.03	1.07	1.14	1.20	1.27	1.33	1.39	1.45	1.51	1.57
0.9	0.8	0.4	1.04	1.07	1.14	1.21	1.27	1.33	1.40	1.46	1.52	1.57
0.9	0.8	0.5	1.04	1.07	1.14	1.20	1.26	1.32	1.38	1.44	1.49	1.55
0.9	0.9	0.1	1.02	1.05	1.10	1.15	1.19	1.24	1.29	1.34	1.39	1.43
0.9	0.9	0.2	1.03	1.07	1.13	1.20	1.26	1.33	1.39	1.46	1.52	1.58
0.9	0.9	0.3	1.04	1.07	1.15	1.22	1.29	1.37	1.44	1.51	1.58	1.65
0.9	0.9	0.4	1.04	1.08	1.16	1.23	1.31	1.38	1.45	1.53	1.60	1.67
0.9	0.9	0.5	1.04	1.08	1.16	1.23	1.31	1.38	1.45	1.52	1.59	1.66
0.9	0.95	0.1	1.03	1.05	1.10	1.15	1.20	1.25	1.30	1.35	1.40	1.45
0.9	0.95	0.2	1.03	1.07	1.14	1.21	1.27	1.34	1.41	1.48	1.55	1.61
0.9	0.95	0.3	1.04	1.08	1.16	1.23	1.31	1.39	1.46	1.54	1.61	1.69
0.9	0.95	0.4	1.04	1.08	1.16	1.25	1.33	1.41	1.49	1.57	1.64	1.72
0.9	0.95	0.5	1.04	1.08	1.17	1.25	1.33	1.42	1.50	1.58	1.65	1.73
0.9	0.99	0.1	1.03	1.05	1.10	1.16	1.21	1.26	1.31	1.37	1.42	1.47
0.9	0.99	0.2	1.04	1.07	1.14	1.21	1.28	1.35	1.42	1.50	1.57	1.64
0.9	0.99	0.3	1.04	1.08	1.16	1.24	1.32	1.40	1.48	1.56	1.64	1.72
0.9	0.99	0.4	1.04	1.09	1.17	1.26	1.34	1.43	1.51	1.60	1.68	1.77
0.9	0.99	0.5	1.04	1.09	1.18	1.27	1.36	1.45	1.53	1.62	1.71	1.80

Table 3-6. Effect of non-differential misclassification of categorical data continued
P_N=the proportion exposed in the non-diseased group

		OR_T	2.0	2.1	2.2	2.3	2.4	2.5	2.6	2.7	2.8	2.9
Specificity	Sensitivity	P_N	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X
0.9	0.1	0.1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.9	0.1	0.2	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.9	0.1	0.3	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.9	0.1	0.4	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.9	0.1	0.5	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.9	0.2	0.1	1.08	1.09	1.10	1.11	1.11	1.12	1.13	1.14	1.14	1.15
0.9	0.2	0.2	1.13	1.14	1.15	1.16	1.17	1.18	1.19	1.20	1.21	1.21
0.9	0.2	0.3	1.15	1.16	1.17	1.18	1.19	1.20	1.21	1.21	1.22	1.23
0.9	0.2	0.4	1.15	1.16	1.17	1.17	1.18	1.19	1.20	1.21	1.21	1.22
0.9	0.2	0.5	1.13	1.14	1.15	1.16	1.17	1.17	1.18	1.19	1.19	1.20
0.9	0.3	0.1	1.16	1.17	1.19	1.20	1.21	1.23	1.24	1.26	1.27	1.28
0.9	0.3	0.2	1.23	1.25	1.27	1.29	1.30	1.32	1.34	1.35	1.37	1.39
0.9	0.3	0.3	1.25	1.27	1.29	1.31	1.32	1.34	1.36	1.37	1.39	1.40
0.9	0.3	0.4	1.24	1.26	1.28	1.29	1.31	1.32	1.34	1.35	1.36	1.37
0.9	0.3	0.5	1.22	1.23	1.25	1.26	1.27	1.28	1.29	1.30	1.31	1.32
0.9	0.4	0.1	1.22	1.24	1.26	1.28	1.30	1.32	1.34	1.36	1.38	1.40
0.9	0.4	0.2	1.31	1.34	1.37	1.39	1.42	1.44	1.47	1.49	1.51	1.53
0.9	0.4	0.3	1.33	1.36	1.39	1.41	1.44	1.46	1.48	1.51	1.53	1.55
0.9	0.4	0.4	1.32	1.34	1.37	1.39	1.41	1.43	1.45	1.47	1.49	1.50
0.9	0.4	0.5	1.29	1.31	1.32	1.34	1.36	1.38	1.39	1.40	1.42	1.43
0.9	0.5	0.1	1.28	1.31	1.34	1.36	1.39	1.41	1.44	1.46	1.49	1.51
0.9	0.5	0.2	1.39	1.42	1.45	1.49	1.52	1.55	1.58	1.61	1.64	1.67
0.9	0.5	0.3	1.41	1.44	1.48	1.51	1.54	1.57	1.60	1.63	1.65	1.68
0.9	0.5	0.4	1.39	1.42	1.45	1.48	1.51	1.53	1.56	1.58	1.60	1.63
0.9	0.5	0.5	1.35	1.38	1.40	1.42	1.44	1.47	1.48	1.50	1.52	1.54
0.9	0.6	0.1	1.34	1.37	1.40	1.43	1.46	1.49	1.53	1.56	1.59	1.62
0.9	0.6	0.2	1.45	1.50	1.54	1.58	1.61	1.65	1.69	1.73	1.76	1.80
0.9	0.6	0.3	1.48	1.52	1.56	1.60	1.64	1.68	1.71	1.75	1.78	1.82
0.9	0.6	0.4	1.47	1.50	1.54	1.57	1.61	1.64	1.67	1.70	1.73	1.76
0.9	0.6	0.5	1.42	1.45	1.48	1.51	1.54	1.56	1.59	1.61	1.64	1.66
0.9	0.7	0.1	1.39	1.43	1.46	1.50	1.54	1.57	1.61	1.64	1.68	1.71
0.9	0.7	0.2	1.52	1.57	1.61	1.66	1.71	1.75	1.80	1.84	1.88	1.93
0.9	0.7	0.3	1.56	1.60	1.65	1.70	1.74	1.79	1.83	1.88	1.92	1.96
0.9	0.7	0.4	1.54	1.59	1.63	1.67	1.72	1.76	1.80	1.83	1.87	1.91
0.9	0.7	0.5	1.50	1.54	1.58	1.61	1.65	1.68	1.71	1.75	1.78	1.81
0.9	0.8	0.1	1.44	1.48	1.52	1.56	1.60	1.65	1.69	1.73	1.77	1.81
0.9	0.8	0.2	1.58	1.64	1.69	1.75	1.80	1.85	1.91	1.96	2.01	2.06
0.9	0.8	0.3	1.63	1.69	1.75	1.80	1.86	1.91	1.97	2.02	2.07	2.12
0.9	0.8	0.4	1.63	1.69	1.74	1.79	1.85	1.90	1.95	1.99	2.04	2.09
0.9	0.8	0.5	1.60	1.65	1.70	1.74	1.79	1.83	1.88	1.92	1.96	2.00
0.9	0.9	0.1	1.48	1.53	1.58	1.62	1.67	1.72	1.77	1.81	1.86	1.91
0.9	0.9	0.2	1.65	1.71	1.77	1.84	1.90	1.96	2.02	2.08	2.14	2.20
0.9	0.9	0.3	1.72	1.78	1.85	1.92	1.99	2.05	2.12	2.18	2.25	2.31
0.9	0.9	0.4	1.74	1.81	1.87	1.94	2.01	2.07	2.14	2.20	2.26	2.32
0.9	0.9	0.5	1.73	1.79	1.86	1.92	1.98	2.04	2.10	2.16	2.22	2.28
0.9	0.95	0.1	1.50	1.55	1.60	1.65	1.70	1.75	1.80	1.85	1.90	1.95
0.9	0.95	0.2	1.68	1.75	1.81	1.88	1.95	2.01	2.08	2.15	2.21	2.28
0.9	0.95	0.3	1.76	1.84	1.91	1.98	2.06	2.13	2.20	2.27	2.35	2.42
0.9	0.95	0.4	1.80	1.88	1.95	2.03	2.10	2.18	2.25	2.33	2.40	2.47
0.9	0.95	0.5	1.81	1.89	1.96	2.04	2.11	2.18	2.26	2.33	2.40	2.47
0.9	0.99	0.1	1.52	1.57	1.63	1.68	1.73	1.78	1.83	1.89	1.94	1.99
0.9	0.99	0.2	1.71	1.78	1.85	1.92	1.99	2.06	2.13	2.20	2.27	2.34
0.9	0.99	0.3	1.80	1.88	1.96	2.04	2.12	2.20	2.28	2.36	2.44	2.51
0.9	0.99	0.4	1.85	1.94	2.02	2.11	2.19	2.28	2.36	2.45	2.53	2.61
0.9	0.99	0.5	1.89	1.98	2.06	2.15	2.24	2.32	2.41	2.50	2.58	2.67

Table 3-6. Effect of non-differential misclassification of categorical data continued
P_N=the proportion exposed in the non-diseased group

		OR_T	3.0	3.1	3.2	3.3	3.4	3.5	3.6	3.7	3.8	3.9
Specificity	Sensitivity	P_N	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X
0.9	0.1	0.1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.9	0.1	0.2	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.9	0.1	0.3	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.9	0.1	0.4	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.9	0.1	0.5	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.9	0.2	0.1	1.16	1.16	1.17	1.18	1.18	1.19	1.19	1.20	1.21	1.21
0.9	0.2	0.2	1.22	1.23	1.24	1.25	1.25	1.26	1.27	1.27	1.28	1.29
0.9	0.2	0.3	1.24	1.25	1.25	1.26	1.27	1.27	1.28	1.29	1.29	1.30
0.9	0.2	0.4	1.23	1.23	1.24	1.25	1.25	1.26	1.26	1.27	1.27	1.28
0.9	0.2	0.5	1.20	1.21	1.21	1.22	1.22	1.23	1.23	1.23	1.24	1.24
0.9	0.3	0.1	1.29	1.31	1.32	1.33	1.34	1.36	1.37	1.38	1.39	1.40
0.9	0.3	0.2	1.40	1.42	1.43	1.44	1.46	1.47	1.49	1.50	1.51	1.52
0.9	0.3	0.3	1.42	1.43	1.44	1.46	1.47	1.48	1.49	1.50	1.51	1.53
0.9	0.3	0.4	1.39	1.40	1.41	1.42	1.43	1.44	1.45	1.46	1.47	1.47
0.9	0.3	0.5	1.33	1.34	1.35	1.36	1.37	1.37	1.38	1.39	1.39	1.40
0.9	0.4	0.1	1.42	1.44	1.46	1.47	1.49	1.51	1.53	1.54	1.56	1.58
0.9	0.4	0.2	1.56	1.58	1.60	1.62	1.64	1.66	1.68	1.70	1.71	1.73
0.9	0.4	0.3	1.57	1.59	1.60	1.62	1.64	1.66	1.67	1.69	1.71	1.72
0.9	0.4	0.4	1.52	1.54	1.55	1.57	1.58	1.59	1.61	1.62	1.63	1.64
0.9	0.4	0.5	1.44	1.46	1.47	1.48	1.49	1.50	1.51	1.52	1.53	1.54
0.9	0.5	0.1	1.54	1.56	1.58	1.61	1.63	1.65	1.68	1.70	1.72	1.74
0.9	0.5	0.2	1.70	1.72	1.75	1.78	1.81	1.83	1.86	1.88	1.91	1.93
0.9	0.5	0.3	1.71	1.73	1.76	1.78	1.80	1.83	1.85	1.87	1.89	1.91
0.9	0.5	0.4	1.65	1.67	1.69	~~1.71~~	1.73	1.74	1.76	1.78	1.80	1.81
0.9	0.5	0.5	1.56	1.57	1.59	1.60	1.62	1.63	1.64	1.65	1.67	1.68
0.9	0.6	0.1	1.65	1.67	1.70	1.73	1.76	1.79	1.82	1.85	1.87	1.90
0.9	0.6	0.2	1.83	1.87	1.90	1.93	1.97	2.00	2.03	2.06	2.09	2.12
0.9	0.6	0.3	1.85	1.88	1.91	1.94	1.97	2.00	2.03	2.06	2.08	2.11
0.9	0.6	0.4	1.78	1.81	1.84	1.86	1.89	1.91	1.93	1.95	1.98	2.00
0.9	0.6	0.5	1.68	1.70	1.72	1.74	1.76	1.78	1.79	1.81	1.83	1.84
0.9	0.7	0.1	1.75	1.78	1.82	1.85	1.89	1.92	1.96	1.99	2.02	2.06
0.9	0.7	0.2	1.97	2.01	2.05	2.09	2.13	2.17	2.21	2.25	2.29	2.33
0.9	0.7	0.3	2.00	2.04	2.08	2.12	2.15	2.19	2.23	2.26	2.30	2.33
0.9	0.7	0.4	1.94	1.98	2.01	2.04	2.07	2.10	2.13	2.16	2.19	2.22
0.9	0.7	0.5	1.83	1.86	1.89	1.91	1.94	1.96	1.98	2.01	2.03	2.05
0.9	0.8	0.1	1.85	1.89	1.93	1.97	2.01	2.05	2.09	2.13	2.17	2.21
0.9	0.8	0.2	2.11	2.16	2.21	2.26	2.31	2.36	2.40	2.45	2.50	2.54
0.9	0.8	0.3	2.17	2.22	2.27	2.32	2.36	2.41	2.46	2.50	2.55	2.59
0.9	0.8	0.4	2.13	2.18	2.22	2.26	2.31	2.35	2.39	2.43	2.47	2.50
0.9	0.8	0.5	2.04	2.07	2.11	2.15	2.18	2.22	2.25	2.28	2.31	2.34
0.9	0.9	0.1	1.95	2.00	2.05	2.09	2.14	2.18	2.23	2.28	2.32	2.37
0.9	0.9	0.2	2.26	2.32	2.38	2.44	2.50	2.56	2.62	2.67	2.73	2.79
0.9	0.9	0.3	2.37	2.44	2.50	2.56	2.62	2.68	2.74	2.80	2.86	2.92
0.9	0.9	0.4	2.39	2.45	2.51	2.56	2.62	2.68	2.74	2.79	2.85	2.90
0.9	0.9	0.5	2.33	2.39	2.44	2.50	2.55	2.60	2.65	2.70	2.75	2.80
0.9	0.95	0.1	2.00	2.05	2.10	2.15	2.20	2.25	2.30	2.35	2.40	2.45
0.9	0.95	0.2	2.34	2.41	2.47	2.54	2.60	2.67	2.73	2.80	2.86	2.92
0.9	0.95	0.3	2.49	2.56	2.63	2.70	2.77	2.84	2.91	2.98	3.05	3.12
0.9	0.95	0.4	2.55	2.62	2.69	2.76	2.83	2.90	2.97	3.04	3.11	3.18
0.9	0.95	0.5	2.54	2.61	2.68	2.75	2.82	2.88	2.95	3.01	3.08	3.14
0.9	0.99	0.1	2.04	2.09	2.15	2.20	2.25	2.30	2.35	2.41	2.46	2.51
0.9	0.99	0.2	2.41	2.48	2.55	2.62	2.69	2.76	2.83	2.90	2.97	3.04
0.9	0.99	0.3	2.59	2.67	2.75	2.83	2.91	2.99	3.07	3.14	3.22	3.30
0.9	0.99	0.4	2.70	2.78	2.86	2.95	3.03	3.11	3.20	3.28	3.36	3.44
0.9	0.99	0.5	2.76	2.84	2.93	3.01	3.10	3.18	3.27	3.35	3.44	3.52

Table 3-6. Effect of non-differential misclassification of categorical data continued

P_N=the proportion exposed in the non-diseased group **OR_X=Observed OR**

		OR_T	1.05	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9
Specificity	Sensitivity	P_N	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X
0.95	0.1	0.1	0.10	1.00	1.01	1.02	1.03	1.03	1.04	1.05	1.06	1.06
0.95	0.1	0.2	0.19	1.01	1.01	1.03	1.04	1.05	1.06	1.08	1.09	1.10
0.95	0.1	0.3	0.29	1.01	1.02	1.03	1.05	1.06	1.08	1.09	1.10	1.11
0.95	0.1	0.4	0.39	1.01	1.02	1.03	1.05	1.06	1.08	1.09	1.10	1.11
0.95	0.1	0.5	0.49	1.01	1.02	1.03	1.05	1.06	1.07	1.08	1.09	1.10
0.95	0.2	0.1	0.10	1.01	1.02	1.04	1.06	1.09	1.11	1.13	1.15	1.17
0.95	0.2	0.2	0.19	1.02	1.03	1.06	1.09	1.12	1.15	1.18	1.20	1.23
0.95	0.2	0.3	0.29	1.02	1.04	1.07	1.10	1.13	1.16	1.19	1.22	1.24
0.95	0.2	0.4	0.39	1.02	1.04	1.07	1.10	1.13	1.16	1.18	1.21	1.23
0.95	0.2	0.5	0.49	1.02	1.03	1.06	1.09	1.12	1.14	1.16	1.18	1.20
0.95	0.3	0.1	0.10	1.02	1.03	1.06	1.10	1.13	1.16	1.19	1.22	1.24
0.95	0.3	0.2	0.19	1.02	1.04	1.09	1.13	1.17	1.21	1.24	1.28	1.32
0.95	0.3	0.3	0.29	1.02	1.05	1.09	1.13	1.18	1.21	1.25	1.29	1.32
0.95	0.3	0.4	0.39	1.02	1.05	1.09	1.13	1.17	1.20	1.24	1.27	1.30
0.95	0.3	0.5	0.49	1.02	1.04	1.08	1.12	1.15	1.18	1.21	1.23	1.26
0.95	0.4	0.1	0.10	1.02	1.04	1.08	1.12	1.16	1.20	1.23	1.27	1.31
0.95	0.4	0.2	0.19	1.03	1.05	1.10	1.15	1.20	1.25	1.29	1.34	1.38
0.95	0.4	0.3	0.29	1.03	1.05	1.11	1.16	1.21	1.25	1.30	1.34	1.38
0.95	0.4	0.4	0.39	1.03	1.05	1.10	1.15	1.20	1.24	1.28	1.32	1.35
0.95	0.4	0.5	0.49	1.02	1.05	1.09	1.13	1.17	1.21	1.24	1.28	1.31
0.95	0.5	0.1	0.10	1.02	1.05	1.09	1.14	1.18	1.23	1.27	1.32	1.36
0.95	0.5	0.2	0.19	1.03	1.06	1.12	1.17	1.23	1.28	1.34	1.39	1.44
0.95	0.5	0.3	0.29	1.03	1.06	1.12	1.18	1.23	1.29	1.34	1.39	1.44
0.95	0.5	0.4	0.39	1.03	1.06	1.12	1.17	1.22	1.27	1.32	1.36	1.40
0.95	0.5	0.5	0.49	1.03	1.05	1.11	1.15	1.20	1.24	1.28	1.32	1.35
0.95	0.6	0.1	0.10	1.03	1.05	1.10	1.16	1.21	1.26	1.31	1.36	1.41
0.95	0.6	0.2	0.19	1.03	1.06	1.13	1.19	1.25	1.31	1.37	1.43	1.49
0.95	0.6	0.3	0.29	1.03	1.07	1.13	1.20	1.26	1.32	1.38	1.43	1.49
0.95	0.6	0.4	0.39	1.03	1.07	1.13	1.19	1.25	1.30	1.36	1.41	1.46
0.95	0.6	0.5	0.49	1.03	1.06	1.12	1.17	1.22	1.27	1.32	1.36	1.41
0.95	0.7	0.1	0.10	1.03	1.06	1.11	1.17	1.23	1.28	1.34	1.39	1.45
0.95	0.7	0.2	0.19	1.04	1.07	1.14	1.21	1.27	1.34	1.40	1.47	1.53
0.95	0.7	0.3	0.29	1.04	1.07	1.14	1.21	1.28	1.35	1.41	1.48	1.54
0.95	0.7	0.4	0.39	1.04	1.07	1.14	1.21	1.27	1.34	1.40	1.46	1.51
0.95	0.7	0.5	0.49	1.03	1.07	1.13	1.19	1.25	1.31	1.36	1.42	1.47
0.95	0.8	0.1	0.10	1.03	1.06	1.12	1.18	1.24	1.31	1.37	1.43	1.48
0.95	0.8	0.2	0.19	1.04	1.07	1.15	1.22	1.29	1.37	1.44	1.51	1.58
0.95	0.8	0.3	0.29	1.04	1.08	1.16	1.23	1.31	1.38	1.45	1.52	1.59
0.95	0.8	0.4	0.39	1.04	1.08	1.15	1.23	1.30	1.37	1.44	1.51	1.58
0.95	0.8	0.5	0.49	1.04	1.08	1.15	1.22	1.29	1.35	1.42	1.48	1.54
0.95	0.9	0.1	0.10	1.03	1.07	1.13	1.20	1.26	1.33	1.39	1.45	1.52
0.95	0.9	0.2	0.19	1.04	1.08	1.16	1.24	1.31	1.39	1.47	1.54	1.62
0.95	0.9	0.3	0.29	1.04	1.08	1.17	1.25	1.33	1.41	1.49	1.57	1.65
0.95	0.9	0.4	0.39	1.04	1.09	1.17	1.25	1.33	1.42	1.50	1.57	1.65
0.95	0.9	0.5	0.49	1.04	1.08	1.17	1.25	1.33	1.41	1.48	1.56	1.63
0.95	0.95	0.1	0.10	1.03	1.07	1.13	1.20	1.27	1.34	1.40	1.47	1.54
0.95	0.95	0.2	0.19	1.04	1.08	1.16	1.24	1.32	1.40	1.48	1.56	1.64
0.95	0.95	0.3	0.29	1.04	1.09	1.17	1.26	1.34	1.43	1.51	1.60	1.68
0.95	0.95	0.4	0.39	1.04	1.09	1.18	1.27	1.35	1.44	1.53	1.61	1.70
0.95	0.95	0.5	0.49	1.04	1.09	1.18	1.27	1.35	1.44	1.52	1.61	1.69
0.95	0.99	0.1	0.10	1.03	1.07	1.14	1.21	1.27	1.34	1.41	1.48	1.55
0.95	0.99	0.2	0.19	1.04	1.08	1.17	1.25	1.33	1.41	1.50	1.58	1.66
0.95	0.99	0.3	0.29	1.04	1.09	1.18	1.27	1.36	1.44	1.53	1.62	1.71
0.95	0.99	0.4	0.39	1.05	1.09	1.18	1.28	1.37	1.46	1.55	1.64	1.73
0.95	0.99	0.5	0.49	1.05	1.09	1.19	1.28	1.38	1.47	1.56	1.65	1.75

Table 3-6. Effect of non-differential misclassification of categorical data continued
P_N=the proportion exposed in the non-diseased group

		OR_T	2.0	2.1	2.2	2.3	2.4	2.5	2.6	2.7	2.8	2.9
Specificity	**Sensitivity**	P_N	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X
0.95	0.1	0.1	1.07	1.08	1.09	1.09	1.10	1.11	1.11	1.12	1.13	1.13
0.95	0.1	0.2	1.11	1.12	1.13	1.14	1.15	1.16	1.17	1.17	1.18	1.19
0.95	0.1	0.3	1.12	1.13	1.14	1.15	1.16	1.17	1.18	1.19	1.20	1.20
0.95	0.1	0.4	1.12	1.13	1.14	1.15	1.16	1.17	1.17	1.18	1.19	1.20
0.95	0.1	0.5	1.11	1.12	1.13	1.14	1.14	1.15	1.16	1.16	1.17	1.17
0.95	0.2	0.1	1.19	1.20	1.22	1.24	1.26	1.28	1.30	1.31	1.33	1.35
0.95	0.2	0.2	1.25	1.28	1.30	1.32	1.35	1.37	1.39	1.41	1.43	1.45
0.95	0.2	0.3	1.27	1.29	1.31	1.33	1.35	1.37	1.39	1.41	1.43	1.45
0.95	0.2	0.4	1.25	1.27	1.29	1.31	1.33	1.34	1.36	1.37	1.39	1.40
0.95	0.2	0.5	1.22	1.24	1.25	1.27	1.28	1.29	1.31	1.32	1.33	1.34
0.95	0.3	0.1	1.27	1.30	1.33	1.36	1.38	1.41	1.44	1.46	1.49	1.51
0.95	0.3	0.2	1.35	1.38	1.42	1.45	1.48	1.51	1.54	1.57	1.60	1.63
0.95	0.3	0.3	1.36	1.39	1.42	1.45	1.48	1.50	1.53	1.55	1.58	1.60
0.95	0.3	0.4	1.33	1.35	1.38	1.40	1.43	1.45	1.47	1.49	1.51	1.53
0.95	0.3	0.5	1.28	1.30	1.32	1.34	1.36	1.38	1.40	1.41	1.43	1.44
0.95	0.4	0.1	1.34	1.38	1.42	1.45	1.49	1.52	1.55	1.59	1.62	1.65
0.95	0.4	0.2	1.43	1.47	1.51	1.55	1.59	1.62	1.66	1.70	1.73	1.77
0.95	0.4	0.3	1.42	1.46	1.50	1.54	1.57	1.61	1.64	1.67	1.70	1.73
0.95	0.4	0.4	1.39	1.42	1.45	1.48	1.51	1.54	1.57	1.59	1.62	1.64
0.95	0.4	0.5	1.33	1.36	1.39	1.41	1.43	1.46	1.48	1.50	1.51	1.53
0.95	0.5	0.1	1.40	1.45	1.49	1.53	1.57	1.61	1.65	1.69	1.73	1.77
0.95	0.5	0.2	1.49	1.54	1.58	1.63	1.68	1.72	1.76	1.81	1.85	1.89
0.95	0.5	0.3	1.48	1.53	1.57	1.62	1.66	1.70	1.74	1.77	1.81	1.85
0.95	0.5	0.4	1.44	1.48	1.52	1.56	1.59	1.63	1.66	1.69	1.72	1.75
0.95	0.5	0.5	1.39	1.42	1.45	1.48	1.51	1.53	1.56	1.58	1.60	1.63
0.95	0.6	0.1	1.46	1.50	1.55	1.60	1.65	1.69	1.74	1.79	1.83	1.88
0.95	0.6	0.2	1.54	1.60	1.65	1.71	1.76	1.81	1.86	1.91	1.96	2.01
0.95	0.6	0.3	1.54	1.59	1.64	1.69	1.74	1.79	1.84	1.88	1.92	1.97
0.95	0.6	0.4	1.50	1.55	1.59	1.64	1.68	1.72	1.76	1.79	1.83	1.86
0.95	0.6	0.5	1.45	1.48	1.52	1.55	1.59	1.62	1.65	1.68	1.71	1.74
0.95	0.7	0.1	1.50	1.56	1.61	1.66	1.72	1.77	1.82	1.87	1.92	1.98
0.95	0.7	0.2	1.59	1.66	1.72	1.78	1.84	1.89	1.95	2.01	2.07	2.12
0.95	0.7	0.3	1.60	1.66	1.72	1.77	1.83	1.89	1.94	1.99	2.04	2.09
0.95	0.7	0.4	1.57	1.62	1.67	1.72	1.77	1.82	1.87	1.91	1.96	2.00
0.95	0.7	0.5	1.51	1.56	1.60	1.65	1.69	1.73	1.76	1.80	1.84	1.87
0.95	0.8	0.1	1.54	1.60	1.66	1.72	1.78	1.84	1.90	1.95	2.01	2.07
0.95	0.8	0.2	1.65	1.71	1.78	1.85	1.92	1.98	2.05	2.11	2.18	2.24
0.95	0.8	0.3	1.66	1.73	1.80	1.86	1.93	1.99	2.05	2.12	2.18	2.24
0.95	0.8	0.4	1.64	1.70	1.77	1.83	1.89	1.94	2.00	2.06	2.11	2.17
0.95	0.8	0.5	1.60	1.65	1.71	1.76	1.81	1.86	1.91	1.96	2.01	2.05
0.95	0.9	0.1	1.58	1.65	1.71	1.78	1.84	1.90	1.97	2.03	2.09	2.16
0.95	0.9	0.2	1.70	1.77	1.85	1.92	2.00	2.07	2.14	2.22	2.29	2.36
0.95	0.9	0.3	1.73	1.81	1.88	1.96	2.04	2.11	2.19	2.26	2.33	2.41
0.95	0.9	0.4	1.73	1.80	1.88	1.95	2.03	2.10	2.17	2.24	2.31	2.38
0.95	0.9	0.5	1.71	1.78	1.85	1.92	1.99	2.05	2.12	2.18	2.25	2.31
0.95	0.95	0.1	1.60	1.67	1.74	1.80	1.87	1.93	2.00	2.07	2.13	2.20
0.95	0.95	0.2	1.72	1.80	1.88	1.96	2.04	2.12	2.20	2.27	2.35	2.43
0.95	0.95	0.3	1.77	1.85	1.93	2.02	2.10	2.18	2.26	2.34	2.42	2.50
0.95	0.95	0.4	1.78	1.86	1.95	2.03	2.11	2.19	2.27	2.36	2.44	2.51
0.95	0.95	0.5	1.78	1.86	1.94	2.02	2.10	2.18	2.26	2.33	2.41	2.49
0.95	0.99	0.1	1.62	1.69	1.75	1.82	1.89	1.96	2.03	2.10	2.16	2.23
0.95	0.99	0.2	1.74	1.83	1.91	1.99	2.07	2.16	2.24	2.32	2.40	2.49
0.95	0.99	0.3	1.80	1.89	1.97	2.06	2.15	2.24	2.33	2.41	2.50	2.59
0.95	0.99	0.4	1.83	1.92	2.01	2.10	2.19	2.28	2.37	2.46	2.55	2.64
0.95	0.99	0.5	1.84	1.93	2.02	2.12	2.21	2.30	2.39	2.48	2.57	2.66

Table 3-6. Effect of non-differential misclassification of categorical data continued
P_N=the proportion exposed in the non-diseased group

		OR_T	3.0	3.1	3.2	3.3	3.4	3.5	3.6	3.7	3.8	3.9
Specificity	Sensitivity	P_N	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X
0.95	0.1	0.1	1.14	1.15	1.15	1.16	1.16	1.17	1.17	1.18	1.19	1.19
0.95	0.1	0.2	1.20	1.21	1.21	1.22	1.23	1.23	1.24	1.25	1.25	1.26
0.95	0.1	0.3	1.21	1.22	1.23	1.23	1.24	1.24	1.25	1.26	1.26	1.27
0.95	0.1	0.4	1.20	1.21	1.21	1.22	1.22	1.23	1.23	1.24	1.24	1.25
0.95	0.1	0.5	1.18	1.18	1.19	1.19	1.20	1.20	1.20	1.21	1.21	1.21
0.95	0.2	0.1	1.36	1.38	1.40	1.41	1.43	1.44	1.46	1.47	1.49	1.50
0.95	0.2	0.2	1.47	1.48	1.50	1.52	1.54	1.55	1.57	1.58	1.60	1.61
0.95	0.2	0.3	1.46	1.48	1.49	1.51	1.52	1.54	1.55	1.56	1.58	1.59
0.95	0.2	0.4	1.42	1.43	1.44	1.45	1.46	1.47	1.48	1.49	1.50	1.51
0.95	0.2	0.5	1.35	1.36	1.37	1.38	1.38	1.39	1.40	1.41	1.41	1.42
0.95	0.3	0.1	1.54	1.56	1.59	1.61	1.64	1.66	1.68	1.70	1.73	1.75
0.95	0.3	0.2	1.65	1.68	1.70	1.73	1.75	1.78	1.80	1.82	1.85	1.87
0.95	0.3	0.3	1.63	1.65	1.67	1.69	1.71	1.73	1.75	1.77	1.79	1.80
0.95	0.3	0.4	1.55	1.57	1.58	1.60	1.62	1.63	1.65	1.66	1.67	1.69
0.95	0.3	0.5	1.46	1.47	1.48	1.49	1.50	1.51	1.53	1.54	1.54	1.55
0.95	0.4	0.1	1.68	1.72	1.75	1.78	1.81	1.84	1.87	1.90	1.93	1.96
0.95	0.4	0.2	1.80	1.83	1.87	1.90	1.93	1.96	1.99	2.02	2.05	2.07
0.95	0.4	0.3	1.76	1.79	1.81	1.84	1.87	1.89	1.92	1.94	1.96	1.98
0.95	0.4	0.4	1.66	1.69	1.71	1.73	1.75	1.77	1.78	1.80	1.82	1.84
0.95	0.4	0.5	1.55	1.57	1.58	1.60	1.61	1.62	1.64	1.65	1.66	1.67
0.95	0.5	0.1	1.81	1.85	1.89	1.92	1.96	2.00	2.03	2.07	2.11	2.14
0.95	0.5	0.2	1.93	1.97	2.01	2.05	2.09	2.12	2.16	2.19	2.23	2.26
0.95	0.5	0.3	1.88	1.92	1.95	1.98	2.01	2.04	2.07	2.10	2.13	2.16
0.95	0.5	0.4	1.78	1.80	1.83	1.85	1.88	1.90	1.92	1.95	1.97	1.99
0.95	0.5	0.5	1.65	1.67	1.69	1.71	1.72	1.74	1.76	1.77	1.79	1.80
0.95	0.6	0.1	1.92	1.97	2.01	2.06	2.10	2.14	2.18	2.23	2.27	2.31
0.95	0.6	0.2	2.05	2.10	2.15	2.19	2.24	2.28	2.32	2.36	2.41	2.45
0.95	0.6	0.3	2.01	2.05	2.09	2.13	2.16	2.20	2.24	2.27	2.31	2.34
0.95	0.6	0.4	1.90	1.93	1.96	1.99	2.02	2.05	2.08	2.11	2.14	2.16
0.95	0.6	0.5	1.76	1.79	1.81	1.83	1.86	1.88	1.90	1.92	1.94	1.96
0.95	0.7	0.1	2.03	2.08	2.13	2.18	2.23	2.28	2.32	2.37	2.42	2.47
0.95	0.7	0.2	2.18	2.23	2.28	2.34	2.39	2.44	2.49	2.54	2.59	2.64
0.95	0.7	0.3	2.14	2.19	2.24	2.29	2.33	2.38	2.42	2.46	2.51	2.55
0.95	0.7	0.4	2.04	2.08	2.12	2.16	2.20	2.23	2.27	2.31	2.34	2.37
0.95	0.7	0.5	1.90	1.94	1.97	2.00	2.03	2.06	2.08	2.11	2.14	2.16
0.95	0.8	0.1	2.12	2.18	2.24	2.29	2.35	2.40	2.46	2.51	2.57	2.62
0.95	0.8	0.2	2.30	2.36	2.43	2.49	2.55	2.61	2.67	2.73	2.78	2.84
0.95	0.8	0.3	2.30	2.36	2.41	2.47	2.53	2.58	2.64	2.69	2.74	2.80
0.95	0.8	0.4	2.22	2.27	2.32	2.37	2.42	2.47	2.51	2.56	2.60	2.65
0.95	0.8	0.5	2.10	2.14	2.18	2.22	2.26	2.30	2.34	2.37	2.41	2.44
0.95	0.9	0.1	2.22	2.28	2.34	2.41	2.47	2.53	2.59	2.65	2.72	2.78
0.95	0.9	0.2	2.44	2.51	2.58	2.65	2.72	2.79	2.86	2.93	3.00	3.07
0.95	0.9	0.3	2.48	2.55	2.62	2.69	2.76	2.83	2.90	2.97	3.04	3.10
0.95	0.9	0.4	2.45	2.52	2.58	2.65	2.71	2.78	2.84	2.90	2.97	3.03
0.95	0.9	0.5	2.37	2.43	2.49	2.55	2.61	2.66	2.72	2.78	2.83	2.88
0.95	0.95	0.1	2.26	2.33	2.40	2.46	2.53	2.59	2.66	2.72	2.79	2.85
0.95	0.95	0.2	2.51	2.59	2.66	2.74	2.82	2.89	2.97	3.05	3.12	3.20
0.95	0.95	0.3	2.58	2.66	2.74	2.82	2.90	2.98	3.06	3.14	3.21	3.29
0.95	0.95	0.4	2.59	2.67	2.75	2.83	2.91	2.98	3.06	3.13	3.21	3.28
0.95	0.95	0.5	2.56	2.64	2.71	2.78	2.86	2.93	3.00	3.07	3.14	3.21
0.95	0.99	0.1	2.30	2.37	2.44	2.51	2.57	2.64	2.71	2.78	2.85	2.92
0.95	0.99	0.2	2.57	2.65	2.73	2.81	2.90	2.98	3.06	3.14	3.22	3.31
0.95	0.99	0.3	2.68	2.76	2.85	2.94	3.03	3.11	3.20	3.29	3.37	3.46
0.95	0.99	0.4	2.73	2.82	2.91	3.00	3.09	3.18	3.27	3.36	3.45	3.53
0.95	0.99	0.5	2.75	2.84	2.93	3.02	3.11	3.20	3.29	3.38	3.47	3.56

Table 3-6. Effect of non-differential misclassification of categorical data continued

P_N=the proportion exposed in the non-diseased group OR_X=Observed OR

		OR_T	1.05	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9
Specificity	**Sensitivity**	P_N	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X
0.99	0.1	0.1	1.02	1.04	1.09	1.13	1.17	1.21	1.25	1.29	1.32	1.36
0.99	0.1	0.2	1.03	1.05	1.10	1.15	1.20	1.24	1.29	1.33	1.37	1.41
0.99	0.1	0.3	1.03	1.05	1.10	1.15	1.19	1.23	1.27	1.31	1.35	1.38
0.99	0.1	0.4	1.02	1.05	1.09	1.13	1.17	1.21	1.24	1.27	1.30	1.33
0.99	0.1	0.5	1.02	1.04	1.08	1.11	1.15	1.17	1.20	1.23	1.25	1.27
0.99	0.2	0.1	1.03	1.06	1.12	1.18	1.24	1.29	1.35	1.40	1.46	1.51
0.99	0.2	0.2	1.03	1.07	1.13	1.19	1.25	1.31	1.36	1.42	1.47	1.52
0.99	0.2	0.3	1.03	1.06	1.12	1.18	1.23	1.28	1.33	1.38	1.42	1.47
0.99	0.2	0.4	1.03	1.06	1.11	1.16	1.20	1.25	1.29	1.33	1.36	1.40
0.99	0.2	0.5	1.02	1.05	1.09	1.13	1.17	1.21	1.24	1.27	1.30	1.32
0.99	0.3	0.1	1.03	1.07	1.14	1.20	1.27	1.34	1.40	1.46	1.53	1.59
0.99	0.3	0.2	1.04	1.07	1.14	1.21	1.28	1.34	1.40	1.46	1.52	1.58
0.99	0.3	[illegible]	1.03	1.07	1.13	1.19	1.25	1.31	1.37	1.42	1.47	1.52
0.99	0.3	0.4	1.03	1.06	1.12	1.17	1.22	1.27	1.32	1.36	1.40	1.44
0.99	0.3	0.5	1.03	1.05	1.10	1.15	1.19	1.23	1.27	1.30	1.33	1.36
0.99	0.4	0.1	1.04	1.07	1.15	1.22	1.29	1.37	1.44	1.50	1.57	1.64
0.99	0.4	0.2	1.04	1.08	1.15	1.22	1.30	1.36	1.43	1.50	1.56	1.63
0.99	0.4	0.3	1.04	1.07	1.14	1.21	1.27	1.33	1.39	1.45	1.51	1.56
0.99	0.4	0.4	1.03	1.07	1.13	1.19	1.24	1.30	1.35	1.39	1.44	1.48
0.99	0.4	0.5	1.03	1.06	1.11	1.16	1.21	1.25	1.29	1.33	1.37	1.40
0.99	0.5	0.1	1.04	1.08	1.16	1.23	1.31	1.39	1.46	1.54	1.61	1.68
0.99	0.5	0.2	1.04	1.08	1.16	1.24	1.31	1.39	1.46	1.53	1.60	1.67
0.99	0.5	0.3	1.04	1.08	1.15	1.22	1.29	1.36	1.42	1.48	1.54	1.60
0.99	0.5	0.4	1.04	1.07	1.14	1.20	1.26	1.32	1.37	1.43	1.48	1.53
0.99	0.5	0.5	1.03	1.06	1.12	1.18	1.23	1.28	1.32	1.37	1.41	1.45
0.99	0.6	0.1	1.04	1.08	1.16	1.24	1.33	1.40	1.48	1.56	1.64	1.72
0.99	0.6	0.2	1.04	1.08	1.17	1.25	1.33	1.40	1.48	1.56	1.63	1.70
0.99	0.6	0.3	1.04	1.08	1.16	1.23	1.31	1.38	1.45	1.52	1.58	1.65
0.99	0.6	0.4	1.04	1.07	1.15	1.22	1.28	1.35	1.41	1.47	1.52	1.58
0.99	0.6	0.5	1.03	1.07	1.13	1.19	1.25	1.30	1.36	1.41	1.45	1.50
0.99	0.7	0.1	1.04	1.09	1.17	1.25	1.34	1.42	1.50	1.58	1.67	1.75
0.99	0.7	0.2	1.04	1.09	1.17	1.26	1.34	1.42	1.50	1.58	1.66	1.74
0.99	0.7	0.3	1.04	1.08	1.17	1.25	1.33	1.40	1.48	1.55	1.62	1.70
0.99	0.7	0.4	1.04	1.08	1.16	1.23	1.30	1.37	1.44	1.51	1.57	1.63
0.99	0.7	0.5	1.04	1.07	1.14	1.21	1.28	1.34	1.40	1.45	1.51	1.56
0.99	0.8	0.1	1.04	1.09	1.17	1.26	1.35	1.43	1.52	1.60	1.69	1.77
0.99	0.8	0.2	1.05	1.09	1.18	1.27	1.35	1.44	1.53	1.61	1.70	1.78
0.99	0.8	0.3	1.04	1.09	1.18	1.26	1.35	1.43	1.51	1.59	1.67	1.75
0.99	0.8	0.4	1.04	1.09	1.17	1.25	1.33	1.41	1.48	1.56	1.63	1.70
0.99	0.8	0.5	1.04	1.08	1.16	1.23	1.31	1.38	1.45	1.51	1.58	1.64
0.99	0.9	0.1	1.04	1.09	1.18	1.27	1.36	1.45	1.54	1.62	1.71	1.80
0.99	0.9	0.2	1.05	1.09	1.19	1.28	1.37	1.46	1.55	1.64	1.73	1.82
0.99	0.9	0.3	1.05	1.09	1.19	1.28	1.37	1.46	1.55	1.63	1.72	1.81
0.99	0.9	0.4	1.05	1.09	1.18	1.27	1.36	1.45	1.53	1.62	1.70	1.78
0.99	0.9	0.5	1.04	1.09	1.18	1.26	1.35	1.43	1.51	1.59	1.67	1.75
0.99	0.95	0.1	1.05	1.09	1.18	1.27	1.36	1.45	1.54	1.63	1.72	1.81
0.99	0.95	0.2	1.05	1.09	1.19	1.28	1.38	1.47	1.56	1.66	1.75	1.84
0.99	0.95	0.3	1.05	1.10	1.19	1.28	1.38	1.47	1.57	1.66	1.75	1.84
0.99	0.95	0.4	1.05	1.09	1.19	1.28	1.38	1.47	1.56	1.65	1.74	1.83
0.99	0.95	0.5	1.05	1.09	1.19	1.28	1.37	1.46	1.55	1.64	1.73	1.81
0.99	0.99	0.1	1.05	1.09	1.18	1.27	1.37	1.46	1.55	1.64	1.73	1.82
0.99	0.99	0.2	1.05	1.10	1.19	1.29	1.38	1.48	1.57	1.67	1.77	1.86
0.99	0.99	0.3	1.05	1.10	1.19	1.29	1.39	1.49	1.58	1.68	1.78	1.87
0.99	0.99	0.4	1.05	1.10	1.20	1.29	1.39	1.49	1.58	1.68	1.78	1.88
0.99	0.99	0.5	1.05	1.10	1.20	1.29	1.39	1.49	1.58	1.68	1.78	1.87

Table 3-6. Effect of non-differential misclassification of categorical data continued
P_N=the proportion exposed in the non-diseased group

		OR_T	2.0	2.1	2.2	2.3	2.4	2.5	2.6	2.7	2.8	2.9
Specificity	**Sensitivity**	P_N	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X
0.99	0.1	0.1	1.40	1.43	1.47	1.50	1.54	1.57	1.61	1.64	1.67	1.70
0.99	0.1	0.2	1.45	1.48	1.52	1.55	1.59	1.62	1.65	1.68	1.71	1.74
0.99	0.1	0.3	1.41	1.45	1.48	1.51	1.53	1.56	1.59	1.61	1.63	1.66
0.99	0.1	0.4	1.36	1.38	1.41	1.43	1.45	1.47	1.49	1.51	1.53	1.54
0.99	0.1	0.5	1.29	1.31	1.33	1.35	1.36	1.38	1.39	1.41	1.42	1.43
0.99	0.2	0.1	1.56	1.61	1.66	1.71	1.76	1.81	1.86	1.91	1.95	2.00
0.99	0.2	0.2	1.57	1.62	1.66	1.71	1.75	1.80	1.84	1.88	1.92	1.96
0.99	0.2	0.3	1.51	1.55	1.59	1.62	1.66	1.69	1.72	1.75	1.79	1.81
0.99	0.2	0.4	1.43	1.46	1.49	1.52	1.55	1.57	1.59	1.62	1.64	1.66
0.99	0.2	0.5	1.35	1.37	1.39	1.42	1.44	1.45	1.47	1.49	1.50	1.52
0.99	0.3	0.1	1.65	1.71	1.77	1.83	1.88	1.94	2.00	2.05	2.11	2.16
0.99	0.3	0.2	1.64	1.69	1.74	1.80	1.85	1.90	1.94	1.99	2.04	2.08
0.99	0.3	0.3	1.56	1.61	1.65	1.69	1.73	1.77	1.81	1.85	1.88	1.92
0.99	0.3	0.4	1.48	1.51	1.55	1.58	1.61	1.64	1.67	1.70	1.72	1.75
0.99	0.3	0.5	1.39	1.42	1.44	1.47	1.49	1.51	1.53	1.55	1.57	1.59
0.99	0.4	0.1	1.71	1.77	1.84	1.90	1.97	2.03	2.09	2.16	2.22	2.28
0.99	0.4	0.2	1.69	1.75	1.81	1.86	1.92	1.97	2.03	2.08	2.13	2.18
0.99	0.4	0.3	1.61	1.66	1.71	1.76	1.80	1.85	1.89	1.93	1.97	2.01
0.99	0.4	0.4	1.53	1.56	1.60	1.64	1.67	1.71	1.74	1.77	1.80	1.83
0.99	0.4	0.5	1.43	1.47	1.49	1.52	1.55	1.57	1.60	1.62	1.64	1.66
0.99	0.5	0.1	1.75	1.83	1.90	1.97	2.04	2.10	2.17	2.24	2.30	2.37
0.99	0.5	0.2	1.73	1.80	1.86	1.92	1.98	2.04	2.10	2.16	2.22	2.27
0.99	0.5	0.3	1.66	1.72	1.77	1.82	1.87	1.92	1.97	2.01	2.06	2.10
0.99	0.5	0.4	1.57	1.62	1.66	1.70	1.74	1.78	1.82	1.86	1.89	1.92
0.99	0.5	0.5	1.48	1.52	1.55	1.58	1.61	1.64	1.67	1.70	1.72	1.75
0.99	0.6	0.1	1.79	1.87	1.94	2.02	2.09	2.16	2.24	2.31	2.38	2.45
0.99	0.6	0.2	1.77	1.85	1.91	1.98	2.05	2.12	2.18	2.24	2.31	2.37
0.99	0.6	0.3	1.71	1.77	1.83	1.89	1.95	2.00	2.05	2.11	2.16	2.21
0.99	0.6	0.4	1.63	1.68	1.73	1.78	1.82	1.87	1.91	1.95	1.99	2.03
0.99	0.6	0.5	1.54	1.58	1.62	1.66	1.69	1.73	1.76	1.79	1.83	1.85
0.99	0.7	0.1	1.83	1.91	1.99	2.06	2.14	2.22	2.30	2.37	2.45	2.53
0.99	0.7	0.2	1.82	1.89	1.97	2.04	2.12	2.19	2.26	2.33	2.40	2.47
0.99	0.7	0.3	1.76	1.83	1.90	1.96	2.03	2.09	2.15	2.21	2.27	2.33
0.99	0.7	0.4	1.69	1.75	1.81	1.87	1.92	1.97	2.02	2.07	2.12	2.17
0.99	0.7	0.5	1.61	1.66	1.71	1.75	1.80	1.84	1.88	1.92	1.96	1.99
0.99	0.8	0.1	1.86	1.94	2.03	2.11	2.19	2.27	2.36	2.44	2.52	2.60
0.99	0.8	0.2	1.86	1.94	2.03	2.11	2.19	2.27	2.34	2.42	2.50	2.57
0.99	0.8	0.3	1.83	1.90	1.98	2.05	2.12	2.20	2.27	2.34	2.40	2.47
0.99	0.8	0.4	1.77	1.84	1.91	1.97	2.04	2.10	2.16	2.22	2.28	2.34
0.99	0.8	0.5	1.70	1.76	1.82	1.87	1.93	1.98	2.03	2.08	2.13	2.18
0.99	0.9	0.1	1.89	1.98	2.06	2.15	2.24	2.32	2.41	2.50	2.58	2.67
0.99	0.9	0.2	1.91	2.00	2.09	2.18	2.26	2.35	2.44	2.52	2.61	2.69
0.99	0.9	0.3	1.90	1.98	2.07	2.15	2.23	2.32	2.40	2.48	2.56	2.64
0.99	0.9	0.4	1.87	1.95	2.03	2.11	2.18	2.26	2.34	2.41	2.49	2.56
0.99	0.9	0.5	1.82	1.90	1.97	2.04	2.11	2.18	2.25	2.32	2.39	2.45
0.99	0.95	0.1	1.90	1.99	2.08	2.17	2.26	2.35	2.44	2.53	2.62	2.71
0.99	0.95	0.2	1.94	2.03	2.12	2.21	2.30	2.39	2.49	2.58	2.67	2.76
0.99	0.95	0.3	1.94	2.03	2.12	2.21	2.30	2.39	2.48	2.57	2.66	2.74
0.99	0.95	0.4	1.92	2.01	2.10	2.19	2.27	2.36	2.45	2.53	2.62	2.70
0.99	0.95	0.5	1.90	1.98	2.07	2.15	2.23	2.32	2.40	2.48	2.56	2.64
0.99	0.99	0.1	1.91	2.01	2.10	2.19	2.28	2.37	2.46	2.55	2.64	2.74
0.99	0.99	0.2	1.96	2.05	2.15	2.24	2.34	2.43	2.53	2.62	2.72	2.81
0.99	0.99	0.3	1.97	2.06	2.16	2.26	2.35	2.45	2.55	2.64	2.74	2.83
0.99	0.99	0.4	1.97	2.07	2.16	2.26	2.36	2.45	2.55	2.64	2.74	2.84
0.99	0.99	0.5	1.97	2.07	2.16	2.26	2.35	2.45	2.54	2.64	2.73	2.83

Table 3-6. Effect of non-differential misclassification of categorical data continued

P_N=the proportion exposed in the non-diseased group

		OR_T	3.0	3.1	3.2	3.3	3.4	3.5	3.6	3.7	3.8	3.9
Specificity	**Sensitivity**	P_N	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X	OR_X
0.99	0.1	0.1	1.73	1.77	1.80	1.83	1.85	1.88	1.91	1.94	1.97	2.00
0.99	0.1	0.2	1.77	1.80	1.83	1.85	1.88	1.90	1.93	1.95	1.98	2.00
0.99	0.1	0.3	1.68	1.70	1.72	1.74	1.76	1.78	1.80	1.82	1.83	1.85
0.99	0.1	0.4	1.56	1.58	1.59	1.61	1.62	1.63	1.65	1.66	1.67	1.68
0.99	0.1	0.5	1.44	1.45	1.47	1.48	1.48	1.49	1.50	1.51	1.52	1.53
0.99	0.2	0.1	2.04	2.09	2.13	2.17	2.22	2.26	2.30	2.34	2.38	2.42
0.99	0.2	0.2	2.00	2.03	2.07	2.10	2.14	2.17	2.20	2.24	2.27	2.30
0.99	0.2	0.3	1.84	1.87	1.90	1.92	1.95	1.97	1.99	2.02	2.04	2.06
0.99	0.2	0.4	1.68	1.70	1.72	1.74	1.76	1.77	1.79	1.81	1.82	1.83
0.99	0.2	0.5	1.53	1.55	1.56	1.57	1.59	1.60	1.61	1.62	1.63	1.64
0.99	0.3	0.1	2.22	2.27	2.32	2.37	2.42	2.47	2.52	2.57	2.62	2.67
0.99	0.3	0.2	2.13	2.17	2.21	2.25	2.29	2.33	2.37	2.41	2.44	2.48
0.99	0.3	0.3	1.95	1.98	2.01	2.04	2.07	2.10	2.13	2.15	2.18	2.20
0.99	0.3	0.4	1.77	1.79	1.82	1.84	1.86	1.88	1.90	1.91	1.93	1.95
0.99	0.3	0.5	1.61	1.62	1.64	1.65	1.67	1.68	1.69	1.71	1.72	1.73
0.99	0.4	0.1	2.34	2.40	2.46	2.51	2.57	2.63	2.68	2.74	2.79	2.85
0.99	0.4	0.2	2.23	2.28	2.33	2.37	2.42	2.46	2.51	2.55	2.59	2.63
0.99	0.4	0.3	2.05	2.08	2.12	2.15	2.19	2.22	2.25	2.28	2.31	2.34
0.99	0.4	0.4	1.86	1.88	1.91	1.94	1.96	1.98	2.01	2.03	2.05	2.07
0.99	0.4	0.5	1.68	1.70	1.72	1.74	1.75	1.77	1.79	1.80	1.81	1.83
0.99	0.5	0.1	2.44	2.50	2.56	2.63	2.69	2.75	2.81	2.88	2.94	3.00
0.99	0.5	0.2	2.33	2.38	2.44	2.49	2.54	2.59	2.64	2.69	2.73	2.78
0.99	0.5	0.3	2.15	2.19	2.23	2.27	2.31	2.35	2.38	2.42	2.45	2.49
0.99	0.5	0.4	1.96	1.99	2.02	2.05	2.08	2.10	2.13	2.16	2.18	2.20
0.99	0.5	0.5	1.77	1.79	1.82	1.84	1.86	1.88	1.90	1.91	1.93	1.95
0.99	0.6	0.1	2.52	2.59	2.66	2.73	2.80	2.87	2.93	3.00	3.07	3.13
0.99	0.6	0.2	2.43	2.49	2.55	2.61	2.66	2.72	2.78	2.83	2.88	2.94
0.99	0.6	0.3	2.26	2.31	2.35	2.40	2.44	2.49	2.53	2.57	2.61	2.65
0.99	0.6	0.4	2.07	2.11	2.14	2.18	2.21	2.25	2.28	2.31	2.34	2.37
0.99	0.6	0.5	1.88	1.91	1.94	1.96	1.99	2.01	2.03	2.06	2.08	2.10
0.99	0.7	0.1	2.60	2.68	2.75	2.83	2.90	2.97	3.05	3.12	3.19	3.26
0.99	0.7	0.2	2.53	2.60	2.67	2.73	2.80	2.86	2.92	2.99	3.05	3.11
0.99	0.7	0.3	2.39	2.44	2.50	2.55	2.60	2.66	2.71	2.76	2.81	2.85
0.99	0.7	0.4	2.21	2.26	2.30	2.35	2.39	2.43	2.47	2.51	2.54	2.58
0.99	0.7	0.5	2.03	2.06	2.10	2.13	2.16	2.19	2.22	2.25	2.28	2.30
0.99	0.8	0.1	2.68	2.76	2.84	2.92	3.00	3.08	3.16	3.24	3.31	3.39
0.99	0.8	0.2	2.65	2.72	2.80	2.87	2.95	3.02	3.09	3.16	3.23	3.30
0.99	0.8	0.3	2.54	2.60	2.67	2.73	2.80	2.86	2.92	2.98	3.04	3.10
0.99	0.8	0.4	2.39	2.45	2.51	2.56	2.61	2.66	2.71	2.76	2.81	2.86
0.99	0.8	0.5	2.23	2.27	2.32	2.36	2.40	2.44	2.48	2.52	2.56	2.60
0.99	0.9	0.1	2.76	2.84	2.93	3.01	3.10	3.18	3.27	3.35	3.44	3.52
0.99	0.9	0.2	2.78	2.86	2.95	3.03	3.11	3.20	3.28	3.36	3.44	3.52
0.99	0.9	0.3	2.72	2.80	2.88	2.96	3.04	3.11	3.19	3.27	3.34	3.42
0.99	0.9	0.4	2.63	2.71	2.78	2.85	2.92	2.99	3.06	3.12	3.19	3.26
0.99	0.9	0.5	2.52	2.58	2.64	2.70	2.76	2.82	2.88	2.94	3.00	3.06
0.99	0.95	0.1	2.80	2.88	2.97	3.06	3.15	3.24	3.33	3.41	3.50	3.59
0.99	0.95	0.2	2.85	2.94	3.03	3.12	3.21	3.30	3.39	3.47	3.56	3.65
0.99	0.95	0.3	2.83	2.92	3.01	3.09	3.18	3.27	3.35	3.44	3.52	3.61
0.99	0.95	0.4	2.79	2.87	2.95	3.04	3.12	3.20	3.28	3.36	3.44	3.52
0.99	0.95	0.5	2.72	2.80	2.87	2.95	3.03	3.10	3.18	3.25	3.32	3.40
0.99	0.99	0.1	2.83	2.92	3.01	3.10	3.19	3.28	3.37	3.46	3.56	3.65
0.99	0.99	0.2	2.91	3.00	3.10	3.19	3.29	3.38	3.48	3.57	3.67	3.76
0.99	0.99	0.3	2.93	3.02	3.12	3.22	3.31	3.41	3.50	3.60	3.69	3.79
0.99	0.99	0.4	2.93	3.03	3.12	3.22	3.31	3.41	3.50	3.59	3.69	3.78
0.99	0.99	0.5	2.92	3.02	3.11	3.20	3.30	3.39	3.48	3.58	3.67	3.76

Table 3-10. Effects of non-differential misclassification of continuous data from equations

ρ_{XT}	ρ^2_{XT}		OR_Xa	OR_Xb	OR_Xa	OR_Xb	OR_Xa	OR_Xb	OR_Xa	OR_Xb	OR_Xa	OR_Xb
		OR_T	1.05		1.10		1.20		1.30		1.40	
0.100	0.010		1.00	1.00	1.01	1.00	1.02	1.00	1.03	1.00	1.03	1.00
0.200	0.040		1.01	1.00	1.02	1.00	1.04	1.01	1.05	1.01	1.07	1.01
0.300	0.090		1.01	1.00	1.03	1.01	1.06	1.02	1.08	1.02	1.11	1.03
0.400	0.160		1.02	1.01	1.04	1.02	1.08	1.03	1.11	1.04	1.14	1.06
0.500	0.250		1.02	1.01	1.05	1.02	1.10	1.05	1.14	1.07	1.18	1.09
0.600	0.360		1.03	1.02	1.06	1.03	1.12	1.07	1.17	1.10	1.22	1.13
0.700	0.490		1.03	1.02	1.07	1.05	1.14	1.09	1.20	1.14	1.27	1.18
0.800	0.640		1.04	1.03	1.08	1.06	1.16	1.12	1.23	1.18	1.31	1.24
0.900	0.810		1.04	1.04	1.09	1.08	1.18	1.16	1.27	1.24	1.35	1.31
0.950	0.903		1.05	1.05	1.09	1.09	1.19	1.18	1.28	1.27	1.38	1.35
0.980	0.960		1.05	1.05	1.10	1.10	1.20	1.19	1.29	1.29	1.39	1.38
0.990	0.980		1.05	1.05	1.10	1.10	1.20	1.20	1.30	1.29	1.40	1.39
		OR_T	1.50		1.60		1.70		1.80		1.90	
0.100	0.010		1.04	1.00	1.05	1.00	1.05	1.01	1.06	1.01	1.07	1.01
0.200	0.040		1.08	1.02	1.10	1.02	1.11	1.02	1.12	1.02	1.14	1.03
0.300	0.090		1.13	1.04	1.15	1.04	1.17	1.05	1.19	1.05	1.21	1.06
0.400	0.160		1.18	1.07	1.21	1.08	1.24	1.09	1.27	1.10	1.29	1.11
0.500	0.250		1.22	1.11	1.26	1.12	1.30	1.14	1.34	1.16	1.38	1.17
0.600	0.360		1.28	1.16	1.33	1.18	1.37	1.21	1.42	1.24	1.47	1.26
0.700	0.490		1.33	1.22	1.39	1.26	1.45	1.30	1.51	1.33	1.57	1.37
0.800	0.640		1.38	1.30	1.46	1.35	1.53	1.40	1.60	1.46	1.67	1.51
0.900	0.810		1.44	1.39	1.53	1.46	1.61	1.54	1.70	1.61	1.78	1.68
0.950	0.903		1.47	1.44	1.56	1.53	1.66	1.61	1.75	1.70	1.84	1.78
0.980	0.960		1.49	1.48	1.59	1.57	1.68	1.66	1.78	1.76	1.88	1.85
0.990	0.980		1.49	1.49	1.59	1.59	1.69	1.68	1.79	1.78	1.89	1.88
		OR_T	2.00		2.10		2.20		2.30		2.40	
0.100	0.010		1.07	1.01	1.08	1.01	1.08	1.01	1.09	1.01	1.09	1.01
0.200	0.040		1.15	1.03	1.16	1.03	1.17	1.03	1.18	1.03	1.19	1.04
0.300	0.090		1.23	1.06	1.25	1.07	1.27	1.07	1.28	1.08	1.30	1.08
0.400	0.160		1.32	1.12	1.35	1.13	1.37	1.13	1.40	1.14	1.42	1.15
0.500	0.250		1.41	1.19	1.45	1.20	1.48	1.22	1.52	1.23	1.55	1.24
0.600	0.360		1.52	1.28	1.56	1.31	1.60	1.33	1.65	1.35	1.69	1.37
0.700	0.490		1.62	1.40	1.68	1.44	1.74	1.47	1.79	1.50	1.85	1.54
0.800	0.640		1.74	1.56	1.81	1.61	1.88	1.66	1.95	1.70	2.01	1.75
0.900	0.810		1.87	1.75	1.95	1.82	2.03	1.89	2.12	1.96	2.20	2.03
0.950	0.903		1.93	1.87	2.02	1.95	2.11	2.04	2.21	2.12	2.30	2.20
0.980	0.960		1.97	1.95	2.07	2.04	2.17	2.13	2.26	2.23	2.36	2.32
0.990	0.980		1.99	1.97	2.08	2.07	2.18	2.17	2.28	2.26	2.38	2.36

$OR_Xa = OR_T\exp(\rho_{XT})$

$OR_Xb = OR_T\exp(\rho^2_{XT})$

Table 3-10. Effects of non-differential misclassification of continuous data from equations

ρ_{XT}	ρ^2_{XT}		OR_Xa	OR_Xb	OR_Xa	OR_Xb	OR_Xa	OR_Xb	OR_Xa	OR_Xb	OR_Xa	OR_Xb
		OR_T	2.50		2.60		2.70		2.80		2.90	
0.100	0.010		1.10	1.01	1.10	1.01	1.10	1.01	1.11	1.01	1.11	1.01
0.200	0.040		1.20	1.04	1.20	1.04	1.22	1.04	1.23	1.04	1.24	1.04
0.300	0.090		1.32	1.09	1.32	1.09	1.35	1.09	1.36	1.10	1.38	1.10
0.400	0.160		1.44	1.16	1.44	1.16	1.49	1.17	1.51	1.18	1.53	1.19
0.500	0.250		1.58	1.26	1.58	1.26	1.64	1.28	1.67	1.29	1.70	1.30
0.600	0.360		1.73	1.39	1.73	1.39	1.81	1.43	1.85	1.45	1.89	1.47
0.700	0.490		1.90	1.57	1.90	1.57	2.00	1.63	2.06	1.66	2.11	1.68
0.800	0.640		2.08	1.80	2.08	1.80	2.21	1.89	2.28	1.93	2.34	1.98
0.900	0.810		2.28	2.10	2.28	2.10	2.44	2.24	2.53	2.30	2.61	2.37
0.950	0.903		2.39	2.29	2.39	2.29	2.57	2.45	2.66	2.53	2.75	2.61
0.980	0.960		2.45	2.41	2.45	2.41	2.65	2.60	2.74	2.69	2.84	2.78
0.990	0.980		2.48	2.45	2.48	2.45	2.67	2.65	2.77	2.74	2.87	2.84
		OR_T	3.00		3.10		3.20		3.30		3.40	
0.100	0.010		1.12	1.01	1.12	1.01	1.12	1.01	1.13	1.01	1.13	1.01
0.200	0.040		1.25	1.04	1.25	1.05	1.26	1.05	1.27	1.05	1.28	1.05
0.300	0.090		1.39	1.10	1.40	1.11	1.42	1.11	1.43	1.11	1.44	1.12
0.400	0.160		1.55	1.19	1.57	1.20	1.59	1.20	1.61	1.21	1.63	1.22
0.500	0.250		1.73	1.32	1.76	1.33	1.79	1.34	1.82	1.35	1.84	1.36
0.600	0.360		1.93	1.49	1.97	1.50	2.01	1.52	2.05	1.54	2.08	1.55
0.700	0.490		2.16	1.71	2.21	1.74	2.26	1.77	2.31	1.80	2.36	1.82
0.800	0.640		2.41	2.02	2.47	2.06	2.54	2.11	2.60	2.15	2.66	2.19
0.900	0.810		2.69	2.43	2.77	2.50	2.85	2.57	2.93	2.63	3.01	2.69
0.950	0.903		2.84	2.70	2.93	2.78	3.02	2.86	3.11	2.94	3.20	3.02
0.980	0.960		2.93	2.87	3.03	2.96	3.13	3.06	3.22	3.15	3.32	3.24
0.990	0.980		2.97	2.94	3.07	3.03	3.16	3.13	3.26	3.22	3.36	3.32
		OR_T	3.50		3.60		3.70		3.80		3.90	
0.100	0.010		1.13	1.01	1.14	1.01	1.14	1.01	1.14	1.01	1.15	1.01
0.200	0.040		1.28	1.05	1.29	1.05	1.30	1.05	1.31	1.05	1.31	1.06
0.300	0.090		1.46	1.12	1.47	1.12	1.48	1.12	1.49	1.13	1.50	1.13
0.400	0.160		1.65	1.22	1.67	1.23	1.69	1.23	1.71	1.24	1.72	1.24
0.500	0.250		1.87	1.37	1.90	1.38	1.92	1.39	1.95	1.40	1.97	1.41
0.600	0.360		2.12	1.57	2.16	1.59	2.19	1.60	2.23	1.62	2.26	1.63
0.700	0.490		2.40	1.85	2.45	1.87	2.50	1.90	2.55	1.92	2.59	1.95
0.800	0.640		2.72	2.23	2.79	2.27	2.85	2.31	2.91	2.35	2.97	2.39
0.900	0.810		3.09	2.76	3.17	2.82	3.25	2.89	3.33	2.95	3.40	3.01
0.950	0.903		3.29	3.10	3.38	3.18	3.47	3.26	3.55	3.34	3.64	3.42
0.980	0.960		3.41	3.33	3.51	3.42	3.60	3.51	3.70	3.60	3.80	3.70
0.990	0.980		3.46	3.41	3.55	3.51	3.65	3.60	3.75	3.70	3.85	3.80

$OR_Xa = OR_T\exp(\rho_{XT})$

$OR_Xb = OR_T\exp(\rho^2_{XT})$

Appendix for Chapter 4

Raw Data, MDLs
and DCAN correction factor

City A daily data from sampling program, µg/L

Water treatment plant #1 reservoir					**Water treatment plant #2 reservoir**				
Date	TCM	DCAN	DCAA	TCAA	Date	TCM	DCAN	DCAA	TCAA
6-Sep-00	13	1.3	7.6	9.7	6-Sep-00	12	1.2	8.7	10
7-Sep-00	17	1.6	9.3	13	7-Sep-00	18	1.6	9.7	14
8-Sep-00	26	2.4	13	14	8-Sep-00	22	1.8	14	13
9-Sep-00	24	2.0	16	14	9-Sep-00	25	1.8	11	13
10-Sep-00	23	1.8	16	14	10-Sep-00	25	1.9	12	11
11-Sep-00	20	1.8	13	15	11-Sep-00	20	1.5	11	14
12-Sep-00	17	1.7	13	13	12-Sep-00	17	1.5	10	11
13-Sep-00	19	1.6	13	14	13-Sep-00	16	1.6	11	12
14-Sep-00	20	1.8	13	15	14-Sep-00	17	1.5	11	13
15-Sep-00	15	1.2	8.4	12	15-Sep-00	13	1.1	9.1	12
16-Sep-00	16	1.2	8.3	11	16-Sep-00	14	1.1	9.5	13
17-Sep-00	23	1.7	9.4	14	17-Sep-00	22	1.7	9.0	15
18-Sep-00	25	2.2	10	16	18-Sep-00	25	1.8	8.5	18
19-Sep-00	20	1.7	10	14	19-Sep-00				
20-Sep-00	14	1.1	7.4	11	20-Sep-00	20	1.7	6.3	8.5
21-Sep-00	13	0.87	7.9	11	21-Sep-00	18	1.3	6.3	10
22-Sep-00	19	1.4	6.9	11	22-Sep-00	18	1.4	5.9	9.1
23-Sep-00	13	0.89	6.7	10	23-Sep-00	12	0.90	6.2	10
24-Sep-00	11	0.80	4.5	7.8	24-Sep-00	11	0.74	5.5	9.0
25-Sep-00	7.2	0.74	3.7	6.5	25-Sep-00	6.8	<0.7	3.1	6.0
26-Sep-00	6.8	0.73	2.9	5.4	26-Sep-00				
27-Sep-00					27-Sep-00	6.9	<0.7	3.2	5.5
28-Sep-00	16	1.0	3.2	8.7	28-Sep-00	16	0.75	3.3	6.1
29-Sep-00					29-Sep-00	19	1.1	3.4	8.8
30-Sep-00	15	1.0	3.0	8.0	30-Sep-00	22	1.4	3.1	8.9
1-Oct-00	12	0.74	2.9	6.5	1-Oct-00	17	0.74		
2-Oct-00	14	0.74	4.0	8.2	2-Oct-00	11	0.76	2.3	6.2
3-Oct-00	12	0.82	4.4	8.3	3-Oct-00	11	0.71		
4-Oct-00	12	0.81	3.2	7.7	4-Oct-00	7.3	0.77	2.8	5.1
Average:	16	1.3	8.2	11	**Average:**	16	1.3	7.4	11
Std Dev:	5.2	0.49	4.2	3.1	**Std Dev:**	5.4	0.40	3.4	3.4
Max:	26	2.4	16	16	**Max:**	25	1.9	14	18
Min:	6.8	0.73	2.9	5.4	**Min:**	6.8	0.71	2.3	5.1

City A daily data from sampling program, continued, μg/L

A 01

Date	TCM	DCAN	DCAA	TCAA	Duplicates TCM	Duplicates DCAN	Duplicates DCAA	Duplicates TCAA
6-Sep-00	11	1.1	8.0	6.7	11	1.1	7.1	7.0
7-Sep-00	15	1.4	9.7	12				
8-Sep-00	21	1.8	12	16				
9-Sep-00	26	2.2	16	17				
10-Sep-00	23	1.9	14	17	21	1.9	14	16
11-Sep-00	19	1.6	14	12				
12-Sep-00	18	1.5	16	15				
13-Sep-00	16	1.5	14	13				
14-Sep-00	17	1.5	15	15	18	1.6	15	15
15-Sep-00	18	1.5	9.9	11				
16-Sep-00	16	0.93	9.9	13				
17-Sep-00	20	1.5	10	16				
18-Sep-00	18	1.4	11	17	19	1.6	14	19
19-Sep-00	19	1.4	10	15				
20-Sep-00	22	1.9	9.7	15				
21-Sep-00	13	1.0	7.7	11				
22-Sep-00	11	1.0	6.8	7.7	12	0.94	8.2	8.7
23-Sep-00	11	0.85	5.9	10				
24-Sep-00	13	0.94	7.1	12				
25-Sep-00	13	0.94	7.1	11				
26-Sep-00	12	0.87	4.3	8.7	13	0.95	4.0	9.1
27-Sep-00	8.1	0.88	3.6	7.0				
28-Sep-00	16	0.83	3.1	8.3				
29-Sep-00	14	0.74	5.1	8.5				
30-Sep-00	17	0.90	3.9	11	15	2.7	4.9	13
1-Oct-00	17	1.6	4.0	12				
2-Oct-00	17	1.9	4.5	14				
3-Oct-00	15	0.87	4.3	12				
4-Oct-00	9	0.79	3.9	8.8				
Average:	16	1.3	8.7	12				
Std Dev:	4.2	0.42	4.1	3.1				
Max:	26	2.2	16	17				
Min:	8.1	0.74	3.1	6.7				

City A daily data from sampling program, continued, µg/L

A 02					Duplicates			
Date	TCM	DCAN	DCAA	TCAA	TCM	DCAN	DCAA	TCAA
6-Sep-00	11	1.1	6.5	5.2				
7-Sep-00	16	1.4	12	13	16	1.4	12	13
8-Sep-00	28	2.3	11	17				
9-Sep-00	22	1.9	17	16				
10-Sep-00	20	1.8	17	17				
11-Sep-00	20	1.9	14	15	22	1.9	14	15
12-Sep-00	16	1.7	11	10				
13-Sep-00	20	1.8	14	12				
14-Sep-00	23	1.8	13	15				
15-Sep-00	16	1.4	12	13	16	1.5	12	13
16-Sep-00	13	1.1	9.4	12				
17-Sep-00	14	1.2	9.5	12				
18-Sep-00	22	1.6	9	14				
19-Sep-00	17	1.4	6.9	12	17	1.4	8.0	14
20-Sep-00	15	1.6	6.7	8.6				
21-Sep-00	15	1.2	8.3	10				
22-Sep-00	13	1.0	8.0	8.2				
23-Sep-00	14	1.3	7.1	11	14	1.1	8.2	13
24-Sep-00	12	0.98	6.8	11				
25-Sep-00	15	1.0	6.8	10				
26-Sep-00	12	0.94	5.8	11				
27-Sep-00	8.9	0.98	3.9	7.4	8.1	0.90	3.7	7.2
28-Sep-00	17	1.0	2.3	6.4				
29-Sep-00	16	0.94	3.8	7.4				
30-Sep-00	14	0.91	3.9	10				
1-Oct-00	15	1.1	4.8	11	15	0.99	4.7	11
2-Oct-00	16	1.0	4.7	11				
3-Oct-00	14	<0.7	2.7	7.5				
4-Oct-00	14	1.0	4.2	11				
Average:	16	1.3	8.4	11				
Std Dev:	4.0	0.38	4.2	3.0				
Max:	28	2.3	17	17				
Min:	8.9	0.91	2.3	5.2				

City A daily data from sampling program, continued, µg/L

A 03

					Duplicates			
Date	TCM	DCAN	DCAA	TCAA	TCM	DCAN	DCAA	TCAA
6-Sep-00	14	1.5	9.5	6.9				
7-Sep-00	15	1.4	11	11				
8-Sep-00	15	1.5	11	12	17	1.6	12	12
9-Sep-00	18	1.7	16	12				
10-Sep-00	17	1.7	16	13				
11-Sep-00	20	2.1	18	17				
12-Sep-00	23	1.9	17	17	25	1.9	17	17
13-Sep-00	23	2.1	16	12				
14-Sep-00	20	2.3	17	9.8				
15-Sep-00	21	2.0	12	13				
16-Sep-00	19	1.5	10	13	21	1.6	10	14
17-Sep-00	21	1.7	11	15				
18-Sep-00	15	1.3	8.0	13				
19-Sep-00	16	1.4	10	12				
20-Sep-00	18	1.5	7.6	14	17	1.4	9.3	15
21-Sep-00	23	2.0	13	15				
22-Sep-00	21	2.0	14	12				
23-Sep-00	17	1.5	12	15				
24-Sep-00	15	1.4	11	13	15	1.3	10	11
25-Sep-00	14	1.3	7.0	11				
26-Sep-00	9.1	0.81	3.7	7.0				
27-Sep-00	8.6	0.90	2.6	5.2				
28-Sep-00	11	1.0	5.7	8.1	9.3	0.88	4.9	7.1
29-Sep-00	17	1.3	4.5	10				
30-Sep-00	19	1.1	6.0	13				
1-Oct-00	18	1.3	4.6	9.7				
2-Oct-00	19	1.5	4.7	9.4	17	1.2	6.1	11
3-Oct-00	16	0.82	5.9	11				
4-Oct-00	14	0.73	3.7	8.4				
Average:	17	1.5	10	12				
Std Dev:	3.8	0.42	4.6	2.9				
Max:	23	2.3	18	17				
Min:	8.6	0.73	2.6	5.2				

City A daily data from sampling program, continued, µg/L

A 04

					Duplicates			
Date	TCM	DCAN	DCAA	TCAA	TCM	DCAN	DCAA	TCAA
6-Sep-00	13	1.3	9.2	8.9				
7-Sep-00	13	1.3	9.6	10				
8-Sep-00	27	2.2	11	17				
9-Sep-00	24	2.2	16	16	23	2.1	15	16
10-Sep-00	28	2.2	19	17				
11-Sep-00	20	1.8	17	12				
12-Sep-00	15	1.7	13	13				
13-Sep-00	21	1.9	11	17	22	1.9	12	17
14-Sep-00	24	1.9	18	13				
15-Sep-00	21	1.7	14	14				
16-Sep-00	14	1.2	8.6	12				
17-Sep-00	13	1.1	8.9	13	14	1.2	9.4	13
18-Sep-00	26	2.3	11	15				
19-Sep-00	19	1.9	14	14				
20-Sep-00	15	1.4	7.9	13				
21-Sep-00	14	1.3	8.7	11	14	1.3	9.6	11
22-Sep-00	15	1.0	7.7	8.1				
23-Sep-00	10	0.77	5.3	7.7				
24-Sep-00	14	1.1	7.1	11				
25-Sep-00	13	0.98	5.5	9.6	15	1.3	5.8	8.7
26-Sep-00	19	1.8	4.6	9.8				
27-Sep-00	6.5	<0.7	3.1	6.3				
28-Sep-00	6.8	<0.7	3.8	5.4				
29-Sep-00	15	1.0	3.5	7.4	14	0.92	3.8	7.3
30-Sep-00	19	1.0	4.2	8.8				
1-Oct-00	15	0.85	3.1	9.1				
2-Oct-00	16	1.5	2.6	7.0				
3-Oct-00	16	0.98	4.8	10	16	0.91	4.4	9.5
4-Oct-00	14	0.96	5.7	12				
Average:	17	1.5	8.9	11				
Std Dev:	5.5	0.48	4.8	3.3				
Max:	28	2.3	19	17				
Min:	6.5	0.77	2.6	5.4				

City B daily data from sampling program, μg/L

Water treatment plant effluent

Date	TCM	BDCM	DCAA	TCAA
6-Sep-00	19	1.1	11	11
7-Sep-00	21	1.2	11	11
8-Sep-00	19	1.1	11	10
9-Sep-00	20	1.1	11	11
10-Sep-00	22	1.1	12	12
11-Sep-00	19	1.0	10	11
12-Sep-00	19	1.0	10	12
13-Sep-00	19	1.0	9	10
14-Sep-00	19	1.1	10	12
15-Sep-00	19	1.0	10	11
16-Sep-00	21	1.1	10	11
17-Sep-00	19	1.0	10	11
18-Sep-00	19	1.0	10	11
19-Sep-00	16	1.0	11	12
20-Sep-00	20	1.1	10	8.7
21-Sep-00	23	1.2	12	10
22-Sep-00	19	1.0	9.4	8.4
23-Sep-00	19	1.1	7.9	7.4
24-Sep-00	20	1.0	9.9	9.0
25-Sep-00	20	0.9	8.9	8.5
26-Sep-00	13	0.9	9.4	9.0
27-Sep-00	20	1.0	9.3	8.9
28-Sep-00	18	1.0	9.2	8.8
29-Sep-00	21	1.1	8.4	8.3
30-Sep-00	19	1.0	8.9	8.4
1-Oct-00	19	1.0	9.2	8.4
2-Oct-00	19	1.1	9.3	8.6
3-Oct-00	15	1.0	7.4	7.5
4-Oct-00	18	1.1	9.1	9.2
Average:	19	1.0	9.8	9.8
Std dev.:	2.0	0.1	1.0	1.4
Max:	23	1.2	12	12
Min:	13	0.9	7.4	7.4

City B daily data from sampling program, continued, μg/L

B 01					Duplicates			
Date	TCM	BDCM	DCAA	TCAA	TCM	BDCM	DCAA	TCAA
6-Sep-00	25	1.1	11	13	25	1.1	13	15
7-Sep-00								
8-Sep-00	27	1.4	14	18				
9-Sep-00	28	1.3	14	17				
10-Sep-00	28	1.3	13	17	28	1.3	13	17
11-Sep-00	20	1.1	10	13				
12-Sep-00	17	1.2	11	13				
13-Sep-00	20	1.1	11	15				
14-Sep-00	30	1.4	10	17	30	1.4	9.8	16
15-Sep-00	25	1.5	12	15				
16-Sep-00	15	1.0	8.5	10				
17-Sep-00	19	1.2	10	13				
18-Sep-00	27	1.4	12	17	27	1.4	11	17
19-Sep-00	17	1.2	8.6	8.8				
20-Sep-00	22	1.3	9.6	11				
21-Sep-00	22	1.3	8.8	10				
22-Sep-00	26	1.6	14	15	27	1.5	15	15
23-Sep-00	27	1.5	13	14				
24-Sep-00	30	1.8	15	16				
25-Sep-00	30	1.5	12	13				
26-Sep-00	26	1.5	10	13	27	1.5	12	14
27-Sep-00	27	1.5	13	14				
28-Sep-00	27	1.5	14	15				
29-Sep-00	17	1.1	8.7	8.5				
30-Sep-00	26	1.4	11	12	26	1.4	11	11
1-Oct-00	24	1.3	10	11				
2-Oct-00	24	1.3	10	11				
3-Oct-00	25	1.5	10	12				
4-Oct-00	24	1.4	11	12				
Average:	24	1.3	11	13				
Std dev.:	4.2	0.2	1.9	2.6				
Max:	30	1.8	15	18				
Min:	15	1.0	8.5	8.5				

Note: Sept 7, volunteer forgot to take sample

City B daily data from sampling program, continued, µg/L

B 02					Duplicates			
Date	**TCM**	**BDCM**	**DCAA**	**TCAA**	**TCM**	**BDCM**	**DCAA**	**TCAA**
6-Sep-00	32	1.7	15	21				
7-Sep-00	31	1.8	16	23	32	1.9	16	24
8-Sep-00	26	1.4	16	21				
9-Sep-00	31	1.7	15	23				
10-Sep-00	27	1.5	15	19				
11-Sep-00	27	1.5	16	21	28	1.5	16	20
12-Sep-00	29	1.6	14	19				
13-Sep-00	39	1.8	18	25				
14-Sep-00	32	1.7	16	22				
15-Sep-00	30	1.7	16	23	31	1.7	14	20
16-Sep-00	29	1.6	16	21				
17-Sep-00	29	1.5	16	21				
18-Sep-00	29	1.6	14	22				
19-Sep-00	29	1.6	17	25	30	1.7	16	24
20-Sep-00	32	1.7	15	17				
21-Sep-00	33	1.7	16	17				
22-Sep-00	31	1.7	14	16				
23-Sep-00	30	1.7	14	17	31	1.7	17	22
24-Sep-00	29	1.8	13	13				
25-Sep-00	29	1.6	15	19				
26-Sep-00	28	1.3	15	18				
27-Sep-00	30	1.5	14	15	30	1.5	15	15
28-Sep-00	31	1.6	16	20				
29-Sep-00	30	1.6	16	19				
30-Sep-00	32	1.7	15	22				
1-Oct-00	32	1.6	15	16	32	1.6	15	16
2-Oct-00	31	1.6	15	15				
3-Oct-00	29	1.7	15	15				
4-Oct-00	32	1.9	16	19				
Average:	30	1.6	15	19				
Std dev.:	2.4	0.1	1.0	3.2				
Max:	39	1.9	18	25				
Min:	26	1.3	13	13				

City B daily data from sampling program, continued, µg/L

B 03					Duplicates			
Date	TCM	BDCM	DCAA	TCAA	TCM	BDCM	DCAA	TCAA
6-Sep-00	28	1.6	15	18				
7-Sep-00	27	1.2	14	16				
8-Sep-00	29	1.3	14	17	27	1.3	14	18
9-Sep-00	27	1.3	14	17				
10-Sep-00	30	1.4	15	20				
11-Sep-00	29	1.3	14	18				
12-Sep-00	29	1.4	14	21	28	1.5	14	21
13-Sep-00	27	1.3	13	16				
14-Sep-00	29	1.5	14	19				
15-Sep-00	36	1.8	9.1	26				
16-Sep-00	28	1.4	16	15	28	1.5	16	18
17-Sep-00	34	1.8	10	19				
18-Sep-00	29	1.5	14	17				
19-Sep-00	28	1.4	14	14				
20-Sep-00	34	1.8	13	19	35	1.8	13	21
21-Sep-00	28	1.4	13	14				
22-Sep-00	27	1.5	12	13				
23-Sep-00	24	1.3	13	13				
24-Sep-00	30	1.5	15	16	26	1.5		
25-Sep-00	23	1.2	11	12				
26-Sep-00	36	1.8	13	20				
27-Sep-00	35	1.7	17	18				
28-Sep-00	28	1.5	13	14	26	1.5	13	14
29-Sep-00	38	2.1	15	19				
30-Sep-00	26	1.4	11	11				
1-Oct-00	26	1.4	11	12				
2-Oct-00	31	1.5	13	13	27	1.5	13	13
3-Oct-00	39	2.1	16	19				
4-Oct-00	26	1.5	13	14				
Average:	30	1.5	13	17				
Std dev.:	4.1	0.2	1.8	3.3				
Max:	39	2.1	17	26				
Min:	23	1.2	9.1	11				

City B daily data from sampling program, continued, μg/L

B 04						**Duplicates**		
Date	TCM	BDCM	DCAA	TCAA	TCM	BDCM	DCAA	TCAA
6-Sep-00	35	1.9	10	27				
7-Sep-00	32	1.8	12	29				
8-Sep-00	34	1.8	12	29				
9-Sep-00	32	1.7	14	25	30	1.7	15	27
10-Sep-00	32	1.7	15	49				
11-Sep-00	32	1.8	13	24				
12-Sep-00	29	1.6	10	23				
13-Sep-00	30	1.7	13	26	31	1.7	13	26
14-Sep-00	34	1.8	14	27				
15-Sep-00	34	1.8	13	27				
16-Sep-00	35	1.8	11	27				
17-Sep-00	33	1.8	13	27	33	1.7	11	22
18-Sep-00	33	1.8	13	26				
19-Sep-00	34	1.8	13	28				
20-Sep-00	32	1.7	13	20				
21-Sep-00	34	1.8	12	20	34	1.8	13	22
22-Sep-00	31	1.7	9.5	18				
23-Sep-00	33	1.8	9.5	16				
24-Sep-00	32	1.8	11	20				
25-Sep-00	30	1.7	13	21	31	1.7	13	19
26-Sep-00	36	1.8	13	22				
27-Sep-00								
28-Sep-00	34	1.7	8.6	16				
29-Sep-00	34	1.7	9.7	18	34	1.7	10	20
30-Sep-00	37	1.8	12	21				
1-Oct-00	38	1.8	13	22				
2-Oct-00	36	1.8	12	19				
3-Oct-00	32	1.9	13	21	32	1.9	13	21
4-Oct-00	34	1.9	12	18				
Average:	33	1.8	12	24				
Std dev.:	2.0	0.1	1.5	6.4				
Max:	38	1.9	15	49				
Min:	29	1.6	9	16				

Note: Sept 27, volunteer forgot to take sample

City A monthly WTP1, µg/L

Date	WTP1TCM	WTP1BDCM	WTP1DCAA	WTP1TCAA	WTP1THAA
Jan-96					
Feb-96			3.4	3.4	7.6
Mar-96					
Apr-96			2.8	0.8	4.0
May-96			3.4	<0.6	3.4
Jun-96			5.7	1.5	7.2
Jul-96			4.7	1.6	6.3
Aug-96			8.1	1.1	9.2
Sep-96					
Oct-96					
Nov-96			1.4	0.2	1.6
Nov-96			1.4	0.2	1.6
Dec-96			1.7	0.2	1.9
Jan-97	1.3	0.4	0.7	0.3	1.0
Feb-97	4.0	0.4	1.0	0.5	1.5
Mar-97	1.6	0.4	2.0	0.5	2.4
Apr-97	1.8	0.4	2.0	0.6	11
May-97	7.7	0.5	7.1	<1	11
Jun-97	24	1.0			
Jul-97	16	0.7	9.5	2.3	<1
Aug-97	8.3	0.5			
Sep-97	6.4	0.4	2.8	<1	2.8
Oct-97	3.6	0.4	1.9	<1	1.9
Nov-97	3.5	0.4	2.3	<1	2.3
Dec-97	8.0	0.6	9.7	1.3	11
Jan-98	12	0.8			
Feb-98	6.0	0.6	6.7	3.4	10
Mar-98	1.7	0.4			
Apr-98	3.2	0.5	3.8	1.3	5.1
May-98	8.5	0.7			
Jun-98	8.4	0.5	5.6	5.9	11
Jul-98	15	0.6			
Aug-98	10	0.6	3.9	1.2	5.0
Sep-98	6.5	0.5	3.9	1.2	5.0
Oct-98	14	0.8	14	7.6	22
Nov-98	6.9	0.7			
Dec-98	7.6	0.7	9.5	4.5	14

City A monthly WTP1, continued, μg/L

Date	WTP1TCM	WTP1BDCM	WTP1DCAA	WTP1TCAA	WTP1THAA
Jan-99	8.5	0.7	6.4	3.3	9.7
Feb-99	7.0	0.6	6.7	2.6	9.3
Mar-99	4.8	0.6	4.8	2.0	6.8
Apr-99	5.6	0.7	2.8	1.6	4.4
May-99	21	0.9	12	6.1	18
Jun-99	18	1.0	9.1	7.3	16
Jul-99	24	0.8	12	8.3	20
Aug-99	22	0.9	9.9	7.0	17
Sep-99	14	0.8	11	5.1	16
Oct-99	9.5	0.7	4.3	2.5	6.8
Nov-99	5.5	0.6	4.0	2.1	6.1
Dec-99	7.5	0.7	3.2	3.1	6.3
Jan-00	6.9	0.6	5.7	4.0	9.7
Feb-00	5.7	0.6	3.6	2.3	5.9
Mar-00	4.3	0.6	5.2	4.9	10.1
Apr-00	5.5	0.7			
May-00	15	1.0	3.4	4.6	8.0
Jun-00	26	1.2	9.1	7.3	16
Jul-00	29	1.3	15	12	27
Aug-00	26	0.8	15	12	26
Sep-00	19	0.9	9.1	9.6	19
Oct-00	14	0.7	6.3	5.3	12
Nov-00	9.6	0.6	5.2	4.8	10
Dec-00	11	0.6	4.2	3.8	8.0

City A monthly WTP2, μg/L

Date	WTP2TCM	WTP2BDCM	WTP2DCAA	WTP2TCAA	WTP2THAA
Jan-96					
Feb-96			3.0	2.7	5.7
Mar-96					
Apr-96			2.1	1.8	4.6
May-96			5.6	2.4	8
Jun-96			8.1	3.1	11.3
Jul-96			4.8	2.9	7.8
Aug-96			4.3	2.1	6.4
Sep-96					
Oct-96					
Nov-96			1.7	0.5	2.2
Nov-96			1.3	0.2	2.2
Dec-96			1.6	0.4	1.9
Jan-97	4.4	0.5	0.6	0.3	0.9
Feb-97	2.3	0.4	0.7	0.3	1.0
Mar-97	2.3	0.4	1.3	0.5	3.1
Apr-97	5.5	0.6	5.6	3.4	12
May-97	17	0.8	8.1	1.3	14.9
Jun-97	27	1.0			
Jul-97	30	0.9	9.6	2.5	13.5
Aug-97	23	0.9			
Sep-97	16	0.7	5.1	<1	5.1
Oct-97	7.2	0.5	2.1	<1	2.1
Nov-97	5.2	0.5	2.7	<1	3.8
Dec-97	4.2	0.5			
Jan-98	5.7	0.6			
Feb-98	3.6	0.5	3.8	1.9	5.7
Mar-98	3.4	0.5			
Apr-98	5.9	0.6	3.4	1.5	4.9
May-98	18	1.0			
Jun-98	16	0.8	9.2	5.3	14.5
Jul-98	26	0.9			
Aug-98	19	0.9	5.7	2.3	8.0
Sep-98	10	0.7	5.7	2.3	8.0
Oct-98	11	0.6	9.6	5.4	15
Nov-98	5.3	0.5			
Dec-98	5.6	0.5	3.9	2.4	6.3

City A monthly WTP2, continued, µg/L

Date	WTP2TCM	WTP2BDCM	WTP2DCAA	WTP2TCAA	WTP2THAA
Jan-99	5.9	0.5	5.5	2.3	7.8
Feb-99	5.3	0.5	5.5	2.1	7.6
Mar-99	3.3	0.5	4.0	1.6	5.6
Apr-99	4.7	0.6	1.3	<1	1.3
May-99	17	0.7	9.3	6.0	15
Jun-99	18	0.9	12	11	23
Jul-99	24	0.7	13	8.7	22
Aug-99	22	0.9	8.8	5.9	15
Sep-99	13	0.8	9.4	4.7	14
Oct-99	7.4	0.6	3.8	2.1	5.9
Nov-99	4.1	0.5	3.2	1.6	4.8
Dec-99	7.1	0.6	2.9	2.9	5.8
Jan-00	7.6	0.6	4.8	3.3	8.1
Feb-00	5.7	0.6	5.4	<1	5.4
Mar-00	3.3	0.5	3.9	2.6	6.5
Apr-00	3.8	0.6			
May-00	13	0.9	2.9	4.0	6.9
Jun-00	19	1.0	12	11	23
Jul-00	31	1.3	14	13	27
Aug-00	27	0.8	16	15	31
Sep-00	18	0.8	6.4	5.9	12
Oct-00	9.0	0.6	5.8	4.8	11
Nov-00	6.5	0.5	3.6	3.0	6.6
Dec-00	8.0	0.5	3.7	3.0	6.7

City A monthly TTHM montoring data at Water treatment plant #1 and Water treatment plant #2 and 12 distribution system sites, μg/L

Date	WTP 1	WTP 2	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10	#11	#12
Jan-94	1.0	1.0												
Feb-94	1.0	1.0	1.0	1.0	1.0	1.0		1.0	1.0	1.0	1.0	1.0	1.0	1.0
Mar-94	1.0	1.0	1.0	1.0	1.0	1.0		1.0	1.0	1.0	1.0	1.0	1.0	1.0
Apr-94	1.0	1.0	1.0	1.0	1.0	1.0		1.0	1.0	1.0	1.0	1.0	1.0	1.0
May-94	2.0	1.0	1.0	1.0	1.0	1.0		1.0	1.0	1.0	1.0	1.0	1.0	1.0
Jun-94	2.0	1.0	1.0	1.0	1.0	2.0		1.0	1.0	2.0	1.0	1.0	1.0	1.0
Jul-94	3.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	1.0	2.0	2.0	2.0	2.0	2.0
Aug-94	3.0	1.0	3.0	2.0	2.0	2.0	3.0	3.0	3.0	2.0	2.0	2.0	2.0	2.0
Sep-94	2.0	1.0	4.0	2.0	2.0	1.0	3.0	3.0	3.0	1.0	1.0	1.0	2.0	3.0
Oct-94	1.0	1.0		1.0	2.0	2.0	2.0	3.0	2.0	1.0	2.0	1.0	2.0	2.0
Nov-94	1.0	1.0		2.0	7.0	1.0	1.0	2.0	2.0	1.0	1.0	1.0	4.0	1.0
Dec-94	1.0	1.0		1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Jan-95	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Feb-95	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Mar-95	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Apr-95	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
May-95	1.0	5.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Jun-95	2.0	9.0	2.0	11	8	7	3.0	3.0	3.0	10	10	10	10	1.0
Jul-95	3.0	14	4.0	15	13	14	7.0	6.0	7.0	15	14	15	15	4.0
Aug-95	3.0	16	4.0	17	17	14	8.0	8.0	9.0		17	17	16	17
Sep-95	2.0	13	6.0	19	19	19	14	7.0	7.0	19	19	19	19	4.0
Oct-95	3.0	7.0	2.0	11	12	11	9.0	6.0	4.0	10		11	10	4.0
Nov-95	3.0	4.0	5.0	6.0	7.0	6.0	6.0	4.0	3.0	6.0		6.0	7.0	4.0
Dec-95	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	5.0	4.0
Jan-96	4.0	9.0	4.0	4.0	5.0	5.0	4.0	4.0	4.0	5.0	4.0	4.0	4.0	4.0
Feb-96	2.0	5.0	4.0	9.0	10	9.0	6.0	5.0	4.0	8.0	9.0	9.0	9.0	4.0
Mar-96	2.0	4.0	4.0	6.0	7.0	6.0	5.0	5.0	4.0	6.0	6.0	6.0	9.0	3.0
Apr-96	2.0	7.0	4.0	6.0	3.0	5.0	4.0	4.0	4.0	4.0	5.0	5.0	6.0	4.0
May-96	5.0	17	5.0	15	12	14	7.0	6.0	7.0	16	16	16	16	6.0
Jun-96	8.0	20	10	22	22	20	13	10	12	26	22	24	27	10
Jul-96	16	22					19	15			27	27	29	14
Aug-96	20	22	21	25	22	20			19	26				
Sep-96	4.0	11					21	12			19	15	17	7.0
Oct-96	3.0	6.0	8.0	9	12	8.0			8.0	8.0				
Nov-96	3.0	6.0					5.0	6.0			6.0	6.0	6.0	3.0
Nov-96														
Dec-96	2.0	6.0	6.0	9.0	8.0	8.0			6.0	7.0				

City A monthly TTHM montoring data at Water treatment plant #1 and Water treatment plant #2s and 12 distribution system sites, continued, μg/L

Date	WTP 1	WTP 2	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10	#11	#12
Jan-97	1.5	4.9					6.0	1.0			9.0	9.0	1.0	1.0
Feb-97	4.5	2.6	4.0	4.0	5.0	4.0			4.0	3.0				
Mar-97	1.7	2.6					4.0	4.0			3.0	3.0	3.0	2.0
Apr-97	1.9	6.0	3.0	7.0	5.0	6.0			4.0	8.0				
May-97	8.3	18					8	5			14	15	15	5
Jun-97	25	28	14	30	27	31			16	34				
Jul-97	16	31					33	32			36	33	36	30
Aug-97	8.8	24	11	32	32	24			13	23				
Sep-97	6.7	16					14	16			22	20	20	9
Oct-97	4.0	7.9	12	14	19	10			11					
Nov-97	3.8	5.7					7.0				9.0	7.0	9.0	5.0
Dec-97	8.9	4.7	6.0	6.0	7.0			10	6.0					
Jan-98	13	6.3					8.0	9.0			3.0	5.0	4.0	12
Feb-98	6.5	4.0	14	5.0	4.0				13					
Mar-98	2.0	3.9					8.0	8.0			5.0	4.0	5.0	2.0
Apr-98	3.8	6.6	5.0	6.0	6.0				5.0	4.0				
May-98	9.6	19					9.0	13			21	22	24	10
Jun-98	9.9	17	12	22	22				13	18				
Jul-98	16	27				26	17	15			30	31	31	18
Aug-98	11	20	19	28	28	25			16					
Sep-98	7.0	11					17	17			10	17	17	12

City A monthly TTHM montoring data at Water treatment plant #1 and Water treatment plant #2s and 12 distribution system sites, continued, µg/L

Date	WTP 1	WTP 2	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10	#11	#12
Oct-98	15	11	10	11	13	10			10	12				
Nov-98	7.5	5.8					15	17			15	10	14	15
Dec-98	8.3	6.1	13	7.0	5.0	7.0			11	6.0				
Jan-99	9.2	6.5					12	11			9.0	9.0	9.0	12
Feb-99	7.6	5.8	15	9.0	9.0	10			14	8.0				
Mar-99	5.2	3.8					10	11			7.0	6.0	7.0	8.0
Apr-99	6.2	4.9	7.0	7.0	6.0	6.0			7.0	5.0				
May-99	22	17					12	10			15	17	16	19
Jun-99	19	19	22	20	20	22			21	20				
Jul-99	24	25					23	21			24	25	25	26
Aug-99	23	22	27	23	25	23			24	22				
Sep-99	15	14					28	26			21	20	18	18
Oct-99	10	8.0	17	12	15	12			13	9.0				
Nov-99	6.1	4.6					10	12			7.0	7.0	6.0	9.0
Dec-99	8.2	7.8	11	9.0	8.0	10			11	8.0				
Jan-00	7.4	8.1												
Feb-00	6.3	6.3												
Mar-00	4.8	3.7												
Apr-00	6.0	4.1												
May-00	16	14												
Jun-00	27	20												
Jul-00	22	17												
Aug-00	26	27												
Sep-00	20	18												
Oct-00	14	9.4												
Nov-00	10	6.8												
Dec-00	11	8.2												

City B Monthly monitoring data, Water treatment plant, μg/L

Date	WTPTCM	WTPBDCM	WTPTTHM	WTPDCAA	WTPTCAA	WTPTHAA
Jan-98	8.1	0.92				
Feb-98	6.5	0.80				
Mar-98	5.9	0.73				
Apr-98	7.3	1.0				
May-98	36	1.7				
Jun-98	47	1.8				
Jul-98	51	2.0				
Aug-98	44	2.5				
Sep-98	25	2.2				
Oct-98	18	0.77	16			
Nov-98	13	0.48	12			
Dec-98	9.5	0.49	10			
Jan-99	6.8	0.47	7.3			
Feb-99	3.9	0.34	4.3	2.9	2.5	5.5
Mar-99	4.7	0.39	5.1	2.6	2.8	5.6
Apr-99	7.0	0.39	7.4	4.7	3.6	8.4
May-99	7.6	0.74	8.4	4.3	4.3	8.7
Jun-99	16	1.1	17	8.3	9.8	18
Jul-99	17	1.2	18	7.9	7.9	16
Aug-99	49	1.9	51	27	22	49
Sep-99	39	1.8	41	21	18	39
Oct-99	16	1.3	17	11	7.5	19
Nov-99	11	1.1	12			0
Dec-99	8.3	1.0	9.4	5.8	3.8	9.7
Jan-00	7.9	1.0	8.9	4.8	3.4	8.2
Feb-00	6	0.94	6.9	3.7	2.2	5.9
Mar-00	4.3	0.75	5.1	3.2	1.9	5.1
Apr-00	8.2	0.91	9.1	5.0	4.1	9.1
May-00	18	1.2	20	16	16	32
Jun-00	18	1.3	20	12	9.2	22
Jul-00	17	1.0	18	14	10	24
Aug-00	18	1.1	20	10	9.5	20
Sep-00	18	1.1	19	11	11	22
Oct-00	20	1.1	21	8.9	8.9	18
Nov-00	13	0.54	13	7.7	8.3	16
Dec-00	12	0.85	13			

City B monthly monitoring data at distribution site #1, μg/L

Date	B01TCM	B01BDCM	B01TTHM	B01DCAA	B01TCAA	B01THAA
Oct-98	26	1.2	27	8.0	11	20
Nov-98	18	0.92	19	6.7	11	19
Dec-98	15	0.85	16	6.3	9.7	16
Jan-99	11	0.79	12	5.1	7.8	13
Feb-99			9.9	4.9	6.9	12
Mar-99			9.0	3.9	5.6	9.8
Apr-99			9.9	5.0	5.4	11
May-99			9.3	5.0	6.0	11
Jun-99			22	10	12	22
Jul-99			28	10	12	22
Aug-99	52	1.8	54	32	32	64
Sep-99			51	25	26	51
Oct-99	28	1.8	29	15	13	28
Nov-99			26			0
Dec-99			13	7.4	6.4	14
Jan-00	12	1.4	13	6.9	6.1	13
Feb-00	12	1.6	13	7.8	7.3	15
Mar-00	7.1	1.2	8.3	5.5	5.3	11
Apr-00	9.2	1.2	10	4.9	4.7	9.6
May-00	30	1.7	31	22	27	48
Jun-00	26	1.6	27	15	13	30
Jul-00			39	25	25	50
Aug-00	28	1.4	30	14	16	30
Sep-00	26	1.6	28	14	17	31
Oct-00	32	1.9	34	14	19	33
Nov-00	22	0.86	23	11	14	25
Dec-00	20	1.8	22			

City B monthly monitoring data at distribution site #2, μg/L

Date	B02TCM	B02BDCM	B02TTHM	B02DCAA	B02TCAA	B02THAA
Oct-98	31	1.4	32	17	0.92	29
Nov-98	21	1.0	22	8.5	14	24
Dec-98	17	0.91	18	7.7	13	21
Jan-99	13	0.93	14	5.6	8.6	14
Feb-99			0	4.5	5.8	11
Mar-99			0	4.5	7.5	12
Apr-99			0	8.0	9.2	18
May-99			0	5.9	8.1	14
Jun-99			0	12	15	27
Jul-99			0	12	14	26
Aug-99	61	2.1	63	36	38	75
Sep-99			0	30	36	67
Oct-99	36	2.2	38	19	18	37
Nov-99			0			0
Dec-99			0	8.1	6.8	15
Jan-00	13	1.6	15	8.5	7.7	16
Feb-00	9.5	0.0	11	6.4	5.9	12
Mar-00	7.5	1.2	8.7	6.3	5.8	12
Apr-00	8.5	1.3	9.8	5.9	5.8	12
May-00	31	1.8	33	24	31	54
Jun-00	33	1.9	35	20	18	39
Jul-00			0	20	20	40
Aug-00	32	1.6	33	15	19	34
Sep-00	31	1.7	33	17	23	40
Oct-00	27	1.6		11	15	27
Nov-00	26	1.0	27	12	16	27
Dec-00	19	1.7	21			

City B monthly monitoring data at distribution site #3, µg/L

Date	B03TCM	B03BDCM	B03TTHM	B03DCAA	B03TCAA	B03THAA
Oct-98	36	1.5	37	4.5	15	20
Nov-98	25	1.2	26	6.4	15	23
Dec-98	22	1.1	23	7.0	15	22
Jan-99	16	1.2	18	6.6	12	19
Feb-99			0	5.2	7.5	13
Mar-99			0	4.4	6.6	11
Apr-99			0	6.5	8.1	15
May-99			0	5.6	7.8	14
Jun-99			0	11	14	26
Jul-99			0	10	12	23
Aug-99	74	2.3	76	41	50	91
Sep-99			0	28	36	65
Oct-99	37	2.5	40	21	20	41
Nov-99			0			0
Dec-99			0	9.6	8.2	18
Jan-00	14	1.6	15	8.7	7.7	16
Feb-00	9.4	1.3	11	6.5	5.2	12
Mar-00	7.5	1.3	8.7	5.9	5.4	11
Apr-00	9.4	1.5	11	6.2	5.8	12
May-00	33	1.9	35	22	31	53
Jun-00	30	1.8	32	18	16	36
Jul-00			0	21	16	37
Aug-00	30	1.5	31	12	14	27
Sep-00	32	1.7	33	18	25	42
Oct-00	36	2.0		15	21	36
Nov-00	23	0.92	24	11	15	26
Dec-00	23	2.1	25			

City B monthly monitoring data at distribution site #4, μg/L

Date	B04TCM	B04BDCM	B04TTHM	B04DCAA	B04TCAA	B04THAA
Oct-98	25	1.1	26	6.2	11	18
Nov-98	17	0.84	18	6.5	11	18
Dec-98	16	0.85	16	6.3	9.2	16
Jan-99	11	0.78	12	5.1	7.6	13
Feb-99			0	4.0	4.7	9.0
Mar-99			0	3.3	4.7	8.4
Apr-99			0	5.3	6.0	12
May-99			0	4.4	5.6	10
Jun-99			0	9.8	11	21
Jul-99			0	9.2	11	21
Aug-99	52	1.8	54	31	32	63
Sep-99			0	26	26	52
Oct-99	29	1.8	30	16	13	29
Nov-99			0			0
Dec-99			0	7.4	5.8	13
Jan-00	12	1.4	13	7.4	6.5	14
Feb-00	9.7	1.4	11	6.5	5.2	12
Mar-00	6.7	1.1	7.8	5.0	3.9	8.9
Apr-00	9.3	1.2	11	5.7	5.6	11
May-00	27	1.5	29	19	22	41
Jun-00	25	1.6	27	16	14	30
Jul-00				16	13	30
Aug-00	27	1.4	29	23	15	38
Sep-00	26	1.5	27	14	18	32
Oct-00	27	1.5				
Nov-00	20	0.82	21	10	13	23
Dec-00	17	1.3	18			

City C monthly TTHM monitoring data at 15 distribution system locations, µg/L

Date	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10	#11	#12	#13	#14	#15
Jan-94	<0.5	80	76		91		93	91	90	93					105
Feb-94	<0.5	65	67		88		87	81	81	90					98
Mar-94	<0.5				90		98	106	97	102					116
Apr-94	<0.5				53		62	59	60						91
May-94	<0.5	123	42		13		60	56	65	77					83
Jun-94	<0.5	132	43												
Jul-94	<0.5	146	33		67				104	86					96
Aug-94	<0.5	102	32				76	66	88	79					77
Sep-94	<0.5	91	32		54		67	69	73	74					79
Oct-94	<0.5		13		39		63	72	81	55					61
Nov-94	<0.5	68	25		39		51	53	48	59					72
Dec-94	<0.5	62	53		67		89	83	79	84					108
Jan-95	<0.5	67	49		78		64	70	85	73					82
Feb-95	<0.5	75	68		94		86		84	94					98
Mar-95	<0.5	82	71		83		89	92		94					121
Apr-95	<0.5	76	57		72		94	87	86	95					100
May-95	<0.5	79	26		55		76	63	61	75					65
Jun-95	<0.5	103	30		57		68	72	73	91					92
Jul-95	<0.5	135	43		83		83	91	80	110					103
Aug-95	<0.5	147	28		68		93	95	78	112					121
Sep-95	<0.5	140	56		67		95	97	99	113					112
Oct-95	<0.5	99	12		107		122	110	118	122					122
Nov-95	<0.5	79	71		71		74	87	78	86					95
Dec-95	<0.5	65	66		85		88	69	84	92					80
Jan-96	<0.5	69	64		79		75	80	84	93					97
Feb-96	<0.5	78	76		87		95	96	98	98					102
Mar-96	<0.5	76			89		92	86	87	98					108
Apr-96	<0.5	72	72		86		97	93	90	95					101
May-96	<0.5	91	42		72		86	80	87	84					94
Jun-96	<0.5	125	44		88		100	96	92	107					117
Jul-96	<0.5														
Aug-96	<0.5	28	162		87		133	112	139	120					119
Sep-96	<0.5	140	42		76		134	92	115	123					119
Oct-96	<0.5	134	69		85		123	98	97	123					123
Nov-96	<0.5		18		35		52	47	47	52					54
Dec-96	<0.5	81	84		96		97	105	99	109					106

NS=no sample taken

City C monthly TTHM monitoring data at 15 distribution system locations, continued, μg/L

Date	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10	#11	#12	#13	#14	#15
Jan-97	<0.5	75	85		87		95	94	95	105					108
Feb-97	<0.5	71	79		82		94	96	90	114					110
Mar-97	<0.5						115	100	109	116					118
Apr-97	<0.5	79	62		73		89	90	86	94					104
May-97	<0.5	141	29		81		95	97	93	101					101
Jun-97	<0.5	123	29		65		81	108	88	100					116
Jul-97	<0.5	160	15		67		105	99	103	122					101
Aug-97	<0.5	160	21		81	76	107	92	98	109					111
Sep-97	<0.5		2.6	3.1	80	72	104	96	102	97					85
Oct-97	<0.5	112		54		78	94	92	91	108					104
Nov-97	<0.5	104	46	37	54	56	63	55	63	62					65
Dec-97	<0.5	86	93	82	111	102	122	118	117	127					112
Jan-98	<0.5	88	95		109	112	122	118	115	116					127
Feb-98	<0.5	89	95		111	107	116	120	115	125					127
Mar-98	<0.5			95		105	131	113	117	129					128
Apr-98	<0.5	88	32	32	61	60	80	72	78	82					80
May-98	<0.5			57	163	109	149	103	113	173					110
Jun-98	<0.5		34	60	60	69	75	86	74	92					95
Jul-98	<0.5	181	41	30	107	85	117	98	111	119					117
Aug-98	<0.5			79	121	125	139	142	132	141					150
Sep-98	<0.5			33		61	129	75	100	93					101
Oct-98	<0.5	103	53	52	70	74	84	97	93	97					109
Nov-98	<0.5	96	3.1	3.7	42	38	46	51	44	53	61	50	67	53	76
Dec-98	<0.5			81	85	85	85	88	90	95	93	98	105	100	101
Jan-99	<0.5	91	NS	NS	115	NS	118	115	116	121	122	124	125	131	129
Feb-99	<0.5	93	96	NS	110	111	120	123	117	126	130	123	127	128	NR
Mar-99	<0.5	98	96	98	113	107	125	118	123	130	127	121	135	132	135
Apr-99	<0.5	42	89	29	67	70	82	85	83	85	89	87	89	90	92
May-99	<0.5	NS	NS	28	64	67	82	80	81	91	96	83	100	99	98
Jun-99	<0.5	131	26	NS	71	71	84	99	87	93	110	96	106	110	111
Jul-99	<0.5	187	NS	46	95	96	141	134	139	151	158	149	151	152	156
Aug-99	<0.5	151	NS	49	88	89	123	119	125	137	132	134	136	138	141
Sep-99	<0.5	132	1	NS	NS	56	74	88	86	94	103	88	99	108	106
Oct-99	<0.05	NS	32	NS	NS	50	81	64	64	74	69	89	77	72	74
Nov-99	<0.05	99	35	NS	70	72	74	74	72	79	80	80	82	82	82
Dec-99	<0.05	96	NS	71	78	82	89	84	82	93	113	95	101	104	97

NS=no sample taken

City C monthly TTHM monitoring data at 15 distribution system locations, continued, µg/L

Date	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10	#11	#12	#13	#14	#15
Jan-00	<0.5	107	NS	119	128	128	137	137	135	143	144	147	146	149	150
Feb-00	<0.5	100	NS	106	114	117	131	123	144	143	137	142	144	138	141
Mar-00	<0.5	110	NS	93	115	120	133	126	129	138	138	145	140	141	144
Apr-00	<0.5	95	NS	27	56	56	68	69	68	74	77	79	79	79	78
May-00	<0.5	102	NS	64	84	91	92	94	99	105	108	97	108	115	113
Jun-00	<0.5	120	NS	29	58	54	75	72	70	88	86	84	89	89	89
Jul-00	<0.5	150	NS	29	75	79	113	106	116	122	122	118	128	125	126
Aug-00		153	NS	50	NS	81	149	109	119	127	124	156	135	133	131
Sep-00	<0.5	26	NS	28	NS	49	NS	NS	82	95	92	95	101	95	97
Oct-00	<0.5	NS	NS	24	43	40	62	65	71	70	72	65	87	78	78
Nov-00	<0.5	119	NS	NS	65	63	74	77	78	86	87	84	93	89	96
Dec-00	<0.5	100	NS	129	135	NS	146	137	145	145	147	143	151	183	157

NS=no sample taken

City C monthly THAA monitoring data at 15 distribution system locations, µg/L

Date	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10	#11	#12	#13	#14	#15
Jan-97	1.5	69	49		68		71	88	76	89					119
Feb-97	<1.0	75	44		67		90	93	97	115					78
Mar-97	<1.0						92	81	68						114
Apr-97	<1.0	72	78		96		74	121	111	114					135
May-97	<1.0	131	76		114		115	125	111	146					140
Jun-97	1.5	78	61		117		93	142	107	131					87
Jul-97	<1.0	99	55		104	74	94	101	87	75					87
Aug-97	<1.0	50	9.2		77	65	63	76	68	71					60
Sep-97	<1.0		4.4	11	39	46	48	64	68	60					115
Oct-97	1.2	83		64		89	102	111	101	115					82
Nov-97	<1.0	71	52	54	70	66	81	84	75	78					88
Dec-97		54	52	53	75	71	82	74	72	85					
Jan-98	<1.0	33	20		29	40	52	62	48	43					70
Feb-98	<1.0	71	50		72	67	82	76	73	80					84
Mar-98	<1.0			67		65	90	74	71	90					80
Apr-98	<1.0	83	46	44	67	61	85	61	75	68					88
May-98	<1.0			54	97	85	84	65	79	92					77
Jun-98	<1.0		3.8	52	50	77	56	76	42	71					77
Jul-98	<1.0	139	35	69	119	112	65	64	80	80					76
Aug-98	<1.0			83	124	129	122	156	106	135					136
Sep-98	<1.0			70		96	106	115	115	109					122
Oct-98	1.3	86	64	63	80	82	99	107	104	105					98
Nov-98	<1.0	56	45	33	50	44	48	48	44	53	50	40	59	46	46
Dec-98	<1.0			76	80	86	86	91	82	86	95	87		88	81
Jan-99	<1.0	84			57		80	70	91	71	83	84	63	87	77
Feb-99	<1.0	71	58		57	56	63	72	62	76	70	66	69	71	66
Mar-99	<1.0	68	28	54	68	79	76	76	85	79	84	75	66	81	76
Apr-99	<1.0	45	83	43	50	75	79	82	68	80	85	87	73	89	82
May-99	<1.0			44	62	77	72	69	84	81	92	60	73	74	93
Jun-99	1.2	110	35		16	18	61	99	61	68	88	68	67	90	84
Jul-99	<1.0	133		59	92	91	117	114	103	122	111	121	81	120	112
Aug-99	11.8	152		101	127	124	141	150	138	158	141	159	112	160	155
Sep-99	<1.0	116	9.2			104	86	82	63	94	132	88	80	127	100
Oct-99	1.7		61			82	77	88	78	90	87	76	73	99	88
Nov-99	<1.0	79	47		94	106	90	91	84	98	93	94	72	92	89
Dec-99	<1.0	82		52	76	74	96	88	84	89	102		83	105	101

City C monthly THAA monitoring data at 15 distribution system locations, continued, µg/L

Date	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10	#11	#12	#13	#14	#15
Jan-00	<1.0	75			58	79	76		66	84	93	81	76	77	83
Feb-00	<1.0	85		48	57	80	67	64	69	79	85	74	65	68	83
Mar-00	<1.0	70		52	72	72	78	76	91	91	84	78	83	87	87
Apr-00	<1.0	86		61	67	74	106	81	89	110	99	93	99	97	90
May-00	<1.0	49		44	49	58	50	73	45	59	77	59	62	64	61
Jun-00															
Jul-00															
Aug-00		86		41		51	28	70	74	64	70	66	73	73	41
Sep-00	<1.0	25		34		38		61	44	78	68	64	43		62
Oct-00	<1.0			32	46	35	48	52	47	55	58	47	35	54	49
Nov-00	<1.0	63			59	62	67	70	63	75	73	71	53	63	68
Dec-00	NA	49		48	54		90	71	65	85	82	88	64	83	83

City C monthly DBP monitoring data at sampling locations #1 and #2, µg/L

Date	TCM1	BDCM1	DCAA1	TCAA1	TCM2	BDCM2	DCAA2	TCAA2
Jan-99	<0.5	<0.5	<0.8	<0.8	83	8.1	26	58
Feb-99	<0.5	<0.5	<0.8	<0.8	85	8.2	25	46
Mar-99	<0.5	<0.5	<0.8	<0.8	89	8.4	24	44
Apr-99	<0.5	<0.5	<0.8	1.1	39	2.7	1.3	44
May-99	<0.5	<0.5	<0.8	<0.8				
Jun-99	<0.5	<0.5	<0.8	1.2	120	11	32	78
Jul-99	<0.5	<0.5	<0.8	<0.8	174	14	45	88
Aug-99	<0.5	<0.5	<0.8	<0.8	139	11	47	92
Sep-99	<0.5	<0.5	<0.8	<0.8	121	11	39	78
Oct-99	<0.5	<0.5	<0.8	1.7				
Nov-99	<0.5	<0.5	<0.8	<0.8	92	7.3	27	52
Dec-99	<0.5	<0.5	<0.8	<0.8	88	8.1	28	54
Jan-00	<0.5	<0.5	<1.0	<1.0	98	9.0	26	49
Feb-00	<0.5	<0.5	<1.0	1.5	91	8.7	24	50
Mar-00	<0.5	<0.5	<1.0	<1.0	101	9.4	24	35
Apr-00	<0.5	<0.5	<1.0	<1.0	87	7.5	28	54
May-00	<0.5	<0.5	<1.0	<1.0	93	9.3	19	30
Jun-00	<0.5	<0.5			109	11		
Jul-00	<0.5	<0.5			138	12		
Aug-00					140	13	31	50
Sep-00	<0.5	<0.5	<1.0	1.6	25	1.4	1.8	24
Oct-00	<0.5	<0.5	<1.0	<1.0				
Nov-00	<0.5	<0.5	<1.0	<1.0	106	10	22	41
Dec-00	<0.5	<0.5			92	8.8	15	32

City C monthly DBP monitoring data at sampling locations #3 and #4, µg/L

Date	TCM3	BDCM3	DCAA3	TCAA3	TCM4	BDCM4	DCAA4	TCAA4
Jan-99								
Feb-99	88	8.0	12	47				
Mar-99	89	6.7	2.0	26	89	8.7	5.3	48
Apr-99	82	7.1	29	54	28	1.7	4.4	38
May-99					26	1.7	<0.8	44
Jun-99	25	1.3	<0.8	35				
Jul-99					44	2.0	<0.8	58
Aug-99					47	2.2	2.6	72
Sep-99	1.4	<0.5	3.2	6.0				
Oct-99	30	2.1	1.8	60				
Nov-99	32	2.5	1.1	46				
Dec-99					65	6.4	1.8	50
Jan-00					109	11		
Feb-00					96	10	6.3	38
Mar-00					85	8.4	5.3	41
Apr-00					25	2.4	7.9	53
May-00					59	5.3	<1.0	44
Jun-00					27	2.2		
Jul-00					28	1.4		
Aug-00					47	3.4	2.4	36
Sep-00					26	1.7	2.9	31
Oct-00					23	1.6	<1.0	32
Nov-00								
Dec-00					118	12	14	30

City C monthly DBP monitoring data at sampling locations #5 and #6, µg/L

Date	TCM5	BDCM5	DCAA5	TCAA5	TCM6	BDCM6	DCAA6	TCAA6
Jan-99	103	12	8.5	48				
Feb-99	99	11	9.4	48	99	11	8.7	48
Mar-99	103	10	13	55	97	10	16	64
Apr-99	62	5.3	8.5	42	64	5.7	14	61
May-99	58	5.7	10	52	62	5.8	14	63
Jun-99	65	6.3	<0.8	16	65	6.5	0.8	17
Jul-99	88	7.4	19	72	89	7.5	19	71
Aug-99	81	6.8	19	95	83	6.8	19	92
Sep-99					51	4.8	14	89
Oct-99					45	4.4	13	70
Nov-99	63	6.3	33	61	65	6.5	37	70
Dec-99	71	7.3	17	59	75	7.7	15	59
Jan-00	116	12	11	48	116	12	13	65
Feb-00	102	11	12	41	105	12	15	62
Mar-00	103	11	15	47	109	12	15	47
Apr-00	51	4.6	14	54	51	4.7	17	57
May-00	76	8.5	8.1	41	82	9.3	13	45
Jun-00	52	6.2			48	6.0		
Jul-00	68	6.9			72	7.4		
Aug-00					74	7.7	10	40
Sep-00					45	3.9	5.7	33
Oct-00	39	4.1	6.2	40	36	3.8	4.2	31
Nov-00	59	5.9	15	44	57	5.6	16	47
Dec-00	122	13	17	34				

City C monthly DBP monitoring data at sampling locations #7 and #8, μg/L

Date	TCM7	BDCM7	DCAA7	TCAA7	TCM8	BDCM8	DCAA8	TCAA8
Jan-99	106	12	16	64	104	12	15	56
Feb-99	108	12	13	50	111	12	15	56
Mar-99	113	12	17	58	107	11	17	58
Apr-99	75	6.9	17	62	78	7.4	18	63
May-99	75	7.0	6.2	66	74	6.8	13	56
Jun-99	76	7.7	7.5	53	90	10	19	81
Jul-99	131	11	32	82	124	10	32	81
Aug-99	114	9.0	31	98	110	8.6	34	103
Sep-99	67	7.0	24	62	81	7.0	16	66
Oct-99	75	6.8	14	64	58	5.8	16	72
Nov-99	67	7.2	31	59	67	7.3	31	60
Dec-99	81	8.2	33	63	76	7.9	28	61
Jan-00	124	13	15	62	123	14		
Feb-00	118	13	16	47	110	13	15	47
Mar-00	120	13	21	50	113	13	20	49
Apr-00	61	6.2	20	86	63	6.0	20	59
May-00	83	9.2	5.4	45	85	9.3	13	60
Jun-00	67	7.9			64	7.9		
Jul-00	104	9.1			97	8.9		
Aug-00	138	11	2.0	26	99	10	20	47
Sep-00							13	48
Oct-00	56	5.8	9.2	39	59	6.0	11	42
Nov-00	67	7.2	21	47	69	7.4	21	49
Dec-00	132	14	32	55	124	13	24	43

City C monthly DBP monitoring data at sampling locations #9 and #10, µg/L

Date	TCM9	BDCM9	DCAA9	TCAA9	TCM10	BDCM10	DCAA10	TCAA10
Jan-99	104	12	15	76	109	12	16	53
Feb-99	105	11	14	49	114	12	16	60
Mar-99	112	12	20	65	118	12	19	60
Apr-99	75	7.2	14	54	78	7.3	19	60
May-99	74	7.0	18	66	83	7.8	14	67
Jun-99	80	7.8	17	42	84	8.5	16	50
Jul-99	127	12	23	78	140	11	36	82
Aug-99	116	9.1	29	95	127	10	41	95
Sep-99	78	7.9	19	45	86	8.3	28	66
Oct-99	58	5.9	15	63	67	6.6	16	74
Nov-99	65	7.1	28	56	71	7.6	36	62
Dec-99	74	7.7	24	60	85	8.7	27	62
Jan-00	122	13	15	51	129	14	19	59
Feb-00	129	14	16	50	128	14	21	55
Mar-00	116	13	24	59	124	14	24	56
Apr-00	62	5.9	21	66	67	6.7	25	85
May-00	89	9.4	6.5	39	95	10	14	45
Jun-00	63	7.5			79	9.4		
Jul-00	107	9.4			112	10		
Aug-00	109	10	21	50	117	10	17	43
Sep-00	75	6.8	3.2	41	87	7.3	20	58
Oct-00	64	6.5	10	38	64	6.5	12	43
Nov-00	71	7.1	18	45	77	8.3	25	50
Dec-00	122	23	21	39	132	14	30	51

City C monthly DBP monitoring data at sampling locations #11 and #12, μg/L

Date	TCM1 1	BDCM1 1	DCAA1 1	TCAA1 1	TCM1 2	BDCM1 2	DCAA1 2	TCAA1 2
Jan-99	110	12	19	65	112	12	19	62
Feb-99	117	13	16	55	111	12	14	53
Mar-99	116	12	21	63	110	11	19	56
Apr-99	81	7.7	21	64	80	7.3	21	65
May-99	88	8.1	21	71	76	7.3	11	50
Jun-99	100	9.4	23	62	87	8.8	16	51
Jul-99	146	12	35	74	138	11	37	82
Aug-99	123	9	28	97	124	10	33	102
Sep-99	96	7.9	30	103	81	7.9	24	64
Oct-99	62	6.2	15	72	81	7.4	19	57
Nov-99	72	7.8	33	60	72	7.7	33	61
Dec-99	107	5.6	37	66	86	8.7		
Jan-00	130	14	22	71	133	14	19	55
Feb-00	123	14	19	60	128	14	21	50
Mar-00	124	14	24	53	130	15	24	52
Apr-00	70	6.7	25	71	73	6.5	25	65
May-00	98	10	19	58	88	9	9.1	50
Jun-00	78	8.5			75	9.0		
Jul-00	112	10			108	10		
Aug-00	113	10	22	43	144	12	20	41
Sep-00	85	7.4	18	50	87	7.5	18	47
Oct-00	65	6.7	12	46	59	6.3	9	38
Nov-00	79	8.4	24	49	76	8.1	23	48
Dec-00	133	14	28	48	128	14	32	52

City C monthly DBP monitoring data at sampling locations #13 and #14, μg/L

Date	TCM13	BDCM13	DCAA13	TCAA13	TCM14	BDCM14	DCAA14	TCAA14
Jan-99	113	12	12	51	118	13	20	64
Feb-99	115	12	14	55	115	12	17	54
Mar-99	123	12	15	51	120	12	21	61
Apr-99	82	7.4	16	57	83	7.7	22	66
May-99	91	8.2	16	57	91	8.5	18	56
Jun-99	97	9.1	17	48	100	10	25	62
Jul-99	140	11	16	62	141	11	36	81
Aug-99	127	10	19	79	128	10	37	100
Sep-99	91	8.1	16	65	99	8.3	27	99
Oct-99	70	6.6	8.8	64	65	6.4	18	81
Nov-99	74	7.7	20	52	74	8.0	32	60
Dec-99	92	8.9	24	59	94	9.4	36	69
Jan-00	132	14	18	58	135	15	19	58
Feb-00	129	14	17	48	124	14	20	49
Mar-00	126	14	22	54	127	14	24	54
Apr-00	72	6.9	24	76	72	7.0	25	70
May-00	98	10	10	52	104	11	19	45
Jun-00	80	9.0			79	10		
Jul-00	118	10			114	10		
Aug-00	124	11	23	45	122	11	23	47
Sep-00	93	7.7	4.7	38	87	7.8		
Oct-00	80	7.4	3.8	31	71	7.2	12	41
Nov-00	85	8.4	13	39	81	8.6	21	42
Dec-00	137	14	19	42	165	17	29	50

City C monthly DBP monitoring data at sampling locations #15, μg/L

Date	TCM15	BDCM15	DCAA15	TCAA15
Jan-99	117	13	14	60
Feb-99			14	52
Mar-99	123	13	17	59
Apr-99	84	7.8	20	62
May-99	90	8.4	21	72
Jun-99	102	10	21	61
Jul-99	144	12	33	77
Aug-99	131	10	37	106
Sep-99	98	8.2	22	79
Oct-99	68	6.6	16	72
Nov-99	74	7.9	30	59
Dec-99	88	8.6	34	67
Jan-00	136	15	21	62
Feb-00	126	14	19	60
Mar-00	130	14	24	53
Apr-00	71	6.8	24	64
May-00	102	10	18	43
Jun-00	80	8.7		
Jul-00	115	10		
Aug-00	121	11	4.0	35
Sep-00	89	7.9	16	46
Oct-00	71	7.2	11	38
Nov-00	86	9.3	23	46
Dec-00	143	14	29	50

City A daily TCM monitoring data at water treatment plant #1, µg/L

Date	Maximum	Minimum	Mean	Median	Conv1	Conv2	Conv3	Conv4
Jan-00	9.2	5.0	6.9	7.1	7.6	7.1	8.4	6.6
Feb-00	9.5	3.3	5.7	5.2	5.2	6.2	8.0	7.3
Mar-00	8.6	2.2	4.3	4.0	2.6	3.3	6.1	5.4
Apr-00	9.4	3.7	5.5	5.2	3.7	4.7	3.9	5.0
May-00	25	7.4	15	15	8.1	18	7.4	7.5
Jun-00	33	17	26	26	20	19	27	27
Jul-00	39	18	29	29	32	27	28	26
Aug-00	37	15	26	26	27	30	37	23
Sep-00	29	12	19	19	20	23	16	17
Oct-00	22	9.0	14	13	15	17	12	12
Nov-00	13	5.6	9.6	10	5.6	8.0		9.5
Dec-00	15	7.3	11	11	7.3	14	9.3	9.6
Jan-99	10	7.5	8.5	8.3	8.3	8.1	9.2	8.5
Feb-99	10	5.0	7.0	7.0	7.5	7.5	8.4	7.3
Mar-99	6.5	2.9	4.8	5.0	5.4	4.2	5.0	5.1
Apr-99	14	2.0	5.6	4.5	7.2	4.8	2.6	5.2
May-99	29	14	21	20	17	17	22	18
Jun-99	27	11	18	18	17	21	18	17
Jul-99	33	19	24	23	23	22	21	20
Aug-99	38	14	22	22	22	22	22	22
Sep-99	19	9.4	14	14	12	13	14	14
Oct-99	13	5.2	9.5	10	9.8	10	10	11
Nov-99	6.8	4.5	5.5	5.3	6.3	6.1	4.9	4.5
Dec-99	13	3.2	7.5	7	13	10	3.2	6.2
Jan-98	20	6.8	12	12	15	16	10	9.3
Feb-98	13	0.7	6.2	7.2	12	1.8	8.2	10
Mar-98	2.6	1.0	1.7	1.6	1.6	1.3	1.4	1.4
Apr-98	6.8	1.9	3.2	3.1	2.0	2.0	3.2	3.1
May-98	13	4.6	8.5	8.1	8.0	9.5	8.7	7.9
Jun-98	14	0.6	8.4	8.5	6.6	0.6	7.6	7.0

Conv1=convenience sampling date: 2nd Tuesday of each month
Conv2=convenience sampling date: the 18th of each month
Conv3=convenience sampling date: 1st Wednesday of each month
Conv4=convenience sampling date: the 4th of each month

City A daily TCM monitoring data at water treatment plant #1, continued, µg/L

Date	Maximum	Minimum	Mean	Median	Conv1	Conv2	Conv3	Conv4
Jul-98	20	11	15	15	15	15	15	17
Aug-98	18	3.8	10	11	7.7	8.3	11	11
Sep-98	11	3.2	6.5	6.2	9.5	5.2	7.5	7.5
Oct-98	26	6.0	14	12	11	14	11	6.0
Nov-98	11	4.1	6.9	6.7	6.6	6.6	7.9	8.2
Dec-98	13	4.8	7.6	7.1	7.2	5.8	5.2	5.0
Jan-97	2.4	0.6	1.3	0.8	0.6	1.7	0.8	0.7
Feb-97	7.2	0.6	4.0	4.2	7.2	3.5	3.4	4.2
Mar-97	2.4	0.9	1.6	1.6	1.9	1.7	1.7	1.3
Apr-97	2.6	0.9	1.8	1.8	1.5	1.7	2.2	1.2
May-97	18	0.6	7.7	6.8	7.4	5.0	5.0	3.6
Jun-97	43	14	24	22	21	27	17	17
Jul-97	37	7.0	16	11	13	10	37	33
Aug-97	13	5.0	8.3	9.0	5.0	6.1	11	9.7
Sep-97	9.4	3.4	6.4	6.5	5.6	4.1	4.8	4.8
Oct-97	5.7	0.6	3.6	3.3	3.4	5.2	5.3	5.1
Nov-97	6.9	1.5	3.5	3.2	6.9	2.5	3.1	3.6
Dec-97	13	3.3	8.0	9.2	3.3	9.1	4.1	3.8

Conv1=convenience sampling date: 2nd Tuesday of each month
Conv2=convenience sampling date: the 18th of each month
Conv3=convenience sampling date: 1st Wednesday of each month
Conv4=convenience sampling date: the 4th of each month

City A daily TTHM monitoring data at water treatment plant #1, µg/L

Date	Maximum	Minimum	Mean	Median	Conv1	Conv2	Conv3	Conv4
Jan-00	9.7	5.6	7.4	7.5	8.1	7.6	9.2	7.2
Feb-00	10	3.8	6.3	5.9	5.8	6.9	8.4	7.9
Mar-00	9.0	2.2	4.8	4.5	3.0	3.7	6.7	6.1
Apr-00	10	4.3	6.2	6.0	4.4	5.3	4.3	5.6
May-00	26	8.5	16	16	8.9	19	8.5	8.7
Jun-00	34	18	27	27	21	20	28	28
Jul-00	41	18	30	30	33	29	30	27
Aug-00	38	15	26	27	29	30	38	24
Sep-00	30	12	20	20	21	25	17	18
Oct-00	24	9.5	14	14	15	18	12	12
Nov-00	14	5.9	10	10	5.9	9.1		10
Dec-00	16	7.3	11	11	7.3	15	9.8	11
Jan-99	11	8.1	9.2	9.0	9.0	8.8	9.9	9.3
Feb-99	11	5.7	7.6	7.6	8.0	8.0	9.0	7.9
Mar-99	7.0	1.0	5.2	5.5	6.1	4.4	5.8	5.7
Apr-99	15	2.3	6.2	5.1	8.6	5.5	3.1	5.9
May-99	30	15	22	21	18	18	24	19
Jun-99	28	12	19	19	18	22	19	18
Jul-99	34	19	24	24	23	22	22	21
Aug-99	40	15	23	23	23	22	23	23
Sep-99	20	9.9	15	14	13	13	15	15
Oct-99	14	6.0	10	11	11	11	11	12
Nov-99	7.4	5.1	6.1	6.1	7.1	7	5.6	5.1
Dec-99	13	3.9	8.2	7.4	13	12	3.9	7.1
Jan-98	21	7.4	13	13	16	17	12	10
Feb-98	13	1.0	6.7	8.0	13	2.0	8.8	11
Mar-98	3.0	1.1	2.0	2.0	1.8	1.4	1.6	2.0
Apr-98	8.0	2.2	3.8	3.4	2.4	2.3	3.5	4.0
May-98	14	5.7	9.6	9.1	9.0	11	9.9	8.8

Conv1=convenience sampling date: 2nd Tuesday of each month
Conv2=convenience sampling date: the 18th of each month
Conv3=convenience sampling date: 1st Wednesday of each month
Conv4=convenience sampling date: the 4th of each month

City A daily TTHM monitoring data at water treatment plant #1, continued, µg/L

Date	Maximum	Minimum	Mean	Median	Conv1	Conv2	Conv3	Conv4
Jun-98	15	7.6	9.9	9.8	7.6	11	9.0	8.0
Jul-98	21	12	16	16	16	16	16	18
Aug-98	18	4.7	11	12	8.8	8.6	12	12
Sep-98	12	3.8	7	6.8	10	6.0	8.0	7.9
Oct-98	27	6.6	15	13	12	15	12	6.6
Nov-98	12	4.5	7.5	7.5	7.0	7.4	8.5	8.8
Dec-98	14	5.5	8.3	7.6	7.9	6.4	5.8	5.7
Jan-97	3.0	1.0	1.5	1.0	1.0	1.7	1.0	1.0
Feb-97	8.0	1.0	4.5	4.6	8.0	4.0	3.6	4.6
Mar-97	2.8	1.0	1.7	1.8	2.1	1.9	1.7	1.3
Apr-97	2.7	1.0	1.9	2.0	1.5	2.0	2.2	1.0
May-97	19	3.0	8.3	7.1	7.9	5.0	5.0	3.8
Jun-97	44	15	25	23	22	27	17	17
Jul-97	38	7.0	16	13	14	10	38	34
Aug-97	13	5.0	8.8	9.4	5.6	6.6	12	10
Sep-97	9.6	4.0	6.7	6.6	5.9	4.3	5.0	5.0
Oct-97	6.5	1.9	4.0	3.7	3.7	5.6	6.0	5.9
Nov-97	7.4	1.6	3.8	3.5	7.1	2.7	3.4	3.6
Dec-97	14	3.7	8.9	10	3.7	9.7	5.0	4.2

Conv1=convenience sampling date: 2nd Tuesday of each month
Conv2=convenience sampling date: the 18th of each month
Conv3=convenience sampling date: 1st Wednesday of each month
Conv4=convenience sampling date: the 4th of each month

City A daily TCM monitoring data at water treatment plant #2, μg/L

Date	Maximum	Minimum	Mean	Median	Conv1	Conv2	Conv3	Conv4
Jan-00	10	4.2	7.6	7.7	8.5	7.7	7.0	5.4
Feb-00	8.6	3.6	5.7	5.6	6.4	5.9	8.3	8.0
Mar-00	5.5	1.9	3.3	3.6	2.6	3.7	4.0	5.5
Apr-00	7.3	2.4	3.8	3.6	2.4	3.0	3.7	3.5
May-00	21	6.8	13	14	8.5	16	6.8	8.4
Jun-00	32	13	19	16	14	16	16	14
Jul-00	41	15	31	31	27	31	27	25
Aug-00	37	15	27	27	28	34	37	26
Sep-00	26	11	18	17	21	23	16	17
Oct-00	15	5.3	9.0	8.9	11	10	7.6	7.6
Nov-00	9.9	4.3	6.5	6.6	8.4	7.4	5.9	6.3
Dec-00	12	4.6	8.0	7.6	5.1	7.6	6.0	5.5
Jan-99	7.6	3.4	5.9	6.1	5.7	5.8	3.4	6.3
Feb-99	7.3	2.4	5.3	5.6	6.5	5.6	5.6	2.4
Mar-99	4.7	0.5	3.3	3.4	4.2	3.2	4.0	3.6
Apr-99	13	1.5	4.7	3.5	3.7	3.7	3.2	3.2
May-99	25	13	17	16	16	16	19	16
Jun-99	27	8.8	18	18	15	18	17	18
Jul-99	31	19	24	24	25	22	25	24
Aug-99	33	15	22	21	23	23	25	25
Sep-99	17	9.7	13	13	11	14	14	12
Oct-99	15	3.9	7.8	7.4	8.6	7.5	7.4	9.9
Nov-99	6.1	3.0	4.1	4.1	4.3	4.1	3.0	3.2
Dec-99	13	4.1	7.0	6.8	9.2	13	4.1	5.2
Jan-98	8.9	1.9	5.7	6.0	6.0	6.7	4.0	3.9
Feb-98	5.6	2.5	3.6	3.4	4.8	2.9	2.8	3.3
Mar-98	6.7	1.7	3.4	3.5	4.0	1.9	3.1	3.5
Apr-98	18	1.4	5.9	4.1	1.4	3.7	3.5	3.5
May-98	23	13	18	18	15	17	19	19
Jun-98	20	0.6	16	16	15	19	15	15

Conv1=convenience sampling date: 2nd Tuesday of each month
Conv2=convenience sampling date: the 18th of each month
Conv3=convenience sampling date: 1st Wednesday of each month
Conv4=convenience sampling date: the 4th of each month

City A daily TCM monitoring data at water treatment plant #2, continued, μg/L

Date	Maximum	Minimum	Mean	Median	Conv1	Conv2	Conv3	Conv4
Jul-98	30	20	26	26	24	23	28	28
Aug-98	24	12	19	19	14	14	23	23
Sep-98	15	6.4	10	10	9.3	11	13	13
Oct-98	21	7.5	11	9.7	7.8	11	8.7	8.3
Nov-98	9.1	2.6	5.3	4.8	6.7	4.5	7.2	6.4
Dec-98	12	3.0	5.6	5.3	5.3	4.7	4.1	4.3
Jan-97	8.6	2.2	4.4	3.9	3.5	4.3	6.1	6.8
Feb-97	3.0	0.6	2.3	2.4	2.7	2.0	2.4	2.4
Mar-97	3.9	1.6	2.3	2.1	2.1	2.2	2.5	2.4
Apr-97	11	2.9	5.5	4.5	3.7	5.5	6.0	4.9
May-97	28	9.3	17	17	17	18	15	14
Jun-97	45	19	27	24	22	34	21	21
Jul-97	43	24	30	29	33	28	31	29
Aug-97	35	12	23	24	14	17	26	31
Sep-97	24	6.1	16	16	24	14	14	14
Oct-97	11	0.6	7.2	6.8	6.9	6.5	10	10
Nov-97	7.4	3.4	5.2	5.1	6.5	5.3	6.0	5.7
Dec-97	6.7	2.5	4.2	4.3	4.8	4.0	3.7	3.8

Conv1=convenience sampling date: 2nd Tuesday of each month
Conv2=convenience sampling date: the 18th of each month
Conv3=convenience sampling date: 1st Wednesday of each month
Conv4=convenience sampling date: the 4th of each month

City A daily TTHM monitoring data at water treatment plant #2, μg/L

Date	Maximum	Minimum	Mean	Median	Conv1	Conv2	Conv3	Conv4
Jan-00	11	4.5	8.1	8.2	9.2	8.2	7.7	5.9
Feb-00	9.1	4.0	6.3	6.4	7.1	6.7	8.7	8.6
Mar-00	6.4	1.9	3.7	3.8	3.1	4.1	4.5	6.0
Apr-00	8.3	3.0	4.2	3.9	3.0	3.5	4.1	3.9
May-00	22	7.3	14	15	9.4	17	7.3	9.2
Jun-00	33	14	20	17	15	17	17	15
Jul-00	43	15	32	33	27	33	28	26
Aug-00	38	16	27	28	28	35	38	26
Sep-00	27	11	18	18	22	24	17	18
Oct-00	15	5.9	9.4	9.0	11	11	8.0	8.0
Nov-00	11	4.3	6.8	6.8	9.1	7.7	5.9	6.5
Dec-00	13	4.6	8.2	7.7	5.1	7.6	6.0	5.5
Jan-99	8.2	3.9	6.5	6.6	6.1	6.3	3.9	6.9
Feb-99	7.8	2.7	5.8	6.0	6.9	6.2	6.0	2.7
Mar-99	5.0	2.2	3.8	3.9	4.7	3.2	4.5	4.1
Apr-99	14	1.0	4.9	3.7	1.0	4.0	3.7	3.2
May-99	25	14	17	17	17	16	20	17
Jun-99	27	9.6	19	19	16	19	18	19
Jul-99	32	19	25	25	26	23	26	25
Aug-99	34	16	22	22	24	24	26	26
Sep-99	18	10	14	14	11	15	15	13
Oct-99	15	4.4	8.4	8.1	9.1	8.1	8.1	10
Nov-99	6.6	3.5	4.6	4.6	4.8	4.6	3.5	4.0
Dec-99	14	4.7	7.6	7.3	9.7	14	4.7	5.8
Jan-98	9.6	2.4	6.3	6.6	6.6	7.0	5.0	4.4
Feb-98	6.1	2.9	4.0	4.0	5.3	3.0	3.2	4.0
Mar-98	7.0	2.0	3.9	4.0	4.5	2.3	3.5	4.0
Apr-98	19	1.9	6.6	4.8	1.9	4.6	3.8	4.0
May-98	24	14	19	19	16	18	20	20
Jun-98	21	14	17	17	15	20	16	16

Conv1=convenience sampling date: 2nd Tuesday of each month
Conv2=convenience sampling date: the 18th of each month
Conv3=convenience sampling date: 1st Wednesday of each month
Conv4=convenience sampling date: the 4th of each month

City A daily TTHM monitoring data at water treatment plant #2, continued, µg/L

Date	Maximum	Minimum	Mean	Median	Conv1	Conv2	Conv3	Conv4
Jul-98	31	21	27	27	25	24	29	29
Aug-98	25	13	20	20	15	15	24	24
Sep-98	16	7.1	11	11	9.6	12	13	14
Oct-98	22	8.1	11	11	8.4	12	9.3	9.0
Nov-98	9.5	2.9	5.8	5.2	7.0	5.0	7.7	6.9
Dec-98	12	3.5	6.1	5.8	5.7	5.1	4.4	4.8
Jan-97	9.2	2.8	4.9	4.5	3.9	4.7	7.0	7.5
Feb-97	3.2	1.0	2.6	2.8	3.0	2.0	2.8	2.7
Mar-97	4.3	1.0	2.6	2.6	2.4	2.7	3.0	2.7
Apr-97	12	3.0	6.0	5.0	4.0	6.0	6.5	5.0
May-97	29	10	18	17	18	19	15	15
Jun-97	46	19	28	25	23	35	21	21
Jul-97	44	24	31	30	34	29	32	30
Aug-97	36	12	24	25	15	18	27	32
Sep-97	25	6.7	16	17	25	15	14	14
Oct-97	12	5.1	7.9	7.5	7.5	7.1	11	11
Nov-97	8.0	3.7	5.7	5.6	7.0	6.0	6.4	6.3
Dec-97	7.3	2.7	4.7	4.9	5.8	4.4	4.0	4.3

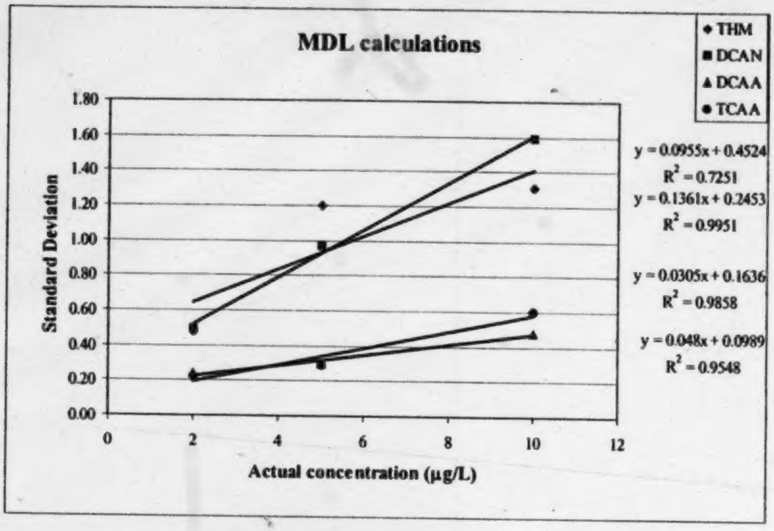
Conv1=convenience sampling date: 2nd Tuesday of each month
Conv2=convenience sampling date: the 18th of each month
Conv3=convenience sampling date: 1st Wednesday of each month
Conv4=convenience sampling date: the 4th of each month

Method detection limit data Edmonton daily data, TCM and DCAN, actual and detected concentrations

Actual conc. µg/L	TCM conc µg/L	DCAN conc µg/L	Actual conc. µg/L	TCM conc µg/L	DCAN conc µg/L	Actual conc. µg/L	TCM conc µg/L	DCAN conc µg/L
2.00	2.27	2.93	5.00	6.94	6.59	10.0	9.90	12.4
2.00	1.93	3.80	5.00	5.48	5.44	10.0	9.72	8.60
2.00	2.09	3.33	5.00	5.07	5.46	10.0	10.2	7.44
2.00	2.05	2.39	5.00	6.61	6.73	10.0	9.18	11.3
2.00	1.83	2.02	5.00	4.67	4.77	10.0	10.9	9.20
2.00	2.15	2.48	5.00	5.30	5.58	10.0	11.5	8.69
2.00	3.12	3.31	5.00	6.52	6.45	10.0	11.7	7.11
2.00	2.27	2.66	5.00	2.99	3.63	10.0	7.87	9.35
2.00	2.63	2.93	5.00	3.94	4.24	10.0	7.29	9.19
2.00	2.81	3.01	5.00	2.42	3.04	10.0	9.34	11.1
2.00	1.57	2.07	5.00	5.13	4.98	10.0	8.05	10.4
2.00	2.47	2.89	5.00	4.50	4.56	10.0	8.52	11.2
2.00	1.66	2.19	5.00	5.49	5.15	10.0	9.61	11.8
2.00	1.73	2.30	5.00	4.89	5.16	10.0	9.74	12.1
2.00	3.22	3.33	5.00	5.50	5.58	10.0	9.26	11.5
2.00	1.91	2.42	5.00	7.26	6.12	10.0	10.1	12.26
2.00	2.22	2.73	5.00	6.91	6.34	10.0	9.36	8.12
2.00	2.17	2.61	5.00	6.24	6.12	10.0	9.41	9.41
2.00	2.36	2.76	5.00	6.57	6.79	10.0	9.58	8.06
2.00	1.51	1.97	5.00	4.13	4.44	10.0	10.1	9.99
2.00	2.00	2.39	5.00	5.48	5.75	10.0	9.56	9.22
2.00	2.16	2.61	5.00	3.41	4.70	10.0	9.05	7.51
2.00	1.63	2.12	5.00	4.74	4.58	10.0	12.6	8.36
2.00	1.52	2.15	5.00	4.74	4.68	10.0	12.7	8.93
2.00	1.56	2.06	5.00	4.16	4.34	10.0	12.0	8.41
2.00	1.70	2.08	5.00	4.46	4.51	10.0	9.73	9.77
2.00	1.49	2.04	5.00	4.01	4.16	10.0	9.47	11.9
ave	2.08	2.58	5.00	4.27	4.23	10.0	9.49	9.33
std dev	0.47	0.49	**ave.**	5.06	5.15	**ave**	9.86	9.74
			std dev	1.22	0.97	**std dev**	1.31	1.59

Method detection limit data Edmonton daily data, DCAA and TCAA, actual and detected concentrations

Actual conc. µg/L	DCAA conc µg/L	TCAA conc µg/L	Actual conc. µg/L	DCAA conc µg/L	TCAA conc µg/L	Actual conc. µg/L	DCAA conc µg/L	TCAA conc µg/L
2.00	1.98	2.03	5.00	5.33	4.78	10.0	9.48	10.3
2.00	2.03	1.87	5.00	4.67	5.12	10.0	10.0	9.68
2.00	1.95	1.66	5.00	4.79	5.09	10.0	9.59	9.55
2.00	1.92	2.05	5.00	5.33	4.74	10.0	10.0	10.1
2.00	2.70	1.44	5.00	5.04	4.83	10.0	9.71	9.28
2.00	1.75	1.68	5.00	4.97	4.98	10.0	10.2	9.33
2.00	1.73	2.06	5.00	4.95	5.15	10.0	9.61	9.08
2.00	2.20	2.17	5.00	5.06	4.80	10.0	10.4	9.74
2.00	1.91	1.91	5.00	4.97	5.34	10.0	9.56	9.45
2.00	1.69	1.87	5.00	5.47	4.94	10.0	9.74	9.37
2.00	1.98	1.81	5.00	4.46	5.60	10.0	9.63	9.22
2.00	1.98	2.15	5.00	4.96	4.90	10.0	10.1	9.79
2.00	2.00	1.99	5.00	4.95	4.93	10.0	9.22	9.06
2.00	1.32	1.47	5.00	5.15	5.27	10.0	9.80	10.2
2.00	2.06	2.37	5.00	4.85	4.98	10.0	9.87	10.0
2.00	1.97	1.89	5.00	4.72	5.17	10.0	9.41	9.72
2.00	1.93	1.85	5.00	5.06	4.69	10.0	10.3	9.88
2.00	1.95	2.18	5.00	5.79	5.14	10.0	10.3	10.9
2.00	2.01	2.16	5.00	5.56	4.00	10.0	9.54	9.81
2.00	1.92	1.94	5.00	4.92	4.70	10.0	9.90	10.8
2.00	2.09	1.90	5.00	5.14	5.07	10.0	10.7	10.3
2.00	1.88	1.91	5.00	5.16	5.10	10.0	10.6	10.9
2.00	2.19	1.66	5.00	4.51	5.08	10.0	10.9	10.9
2.00	1.96	1.80	5.00	4.80	5.03	10.0	10.7	10.8
ave	1.96	1.91	5.00	4.93	4.81	10.0	10.7	9.53
std dev	0.24	0.23	5.00	5.22	4.54	**ave**	10.0	9.91
			5.00	5.31	4.77	**std dev**	0.48	0.60
			5.00	4.93	4.82			
			ave	5.04	4.94			
			std dev	0.30	0.29			

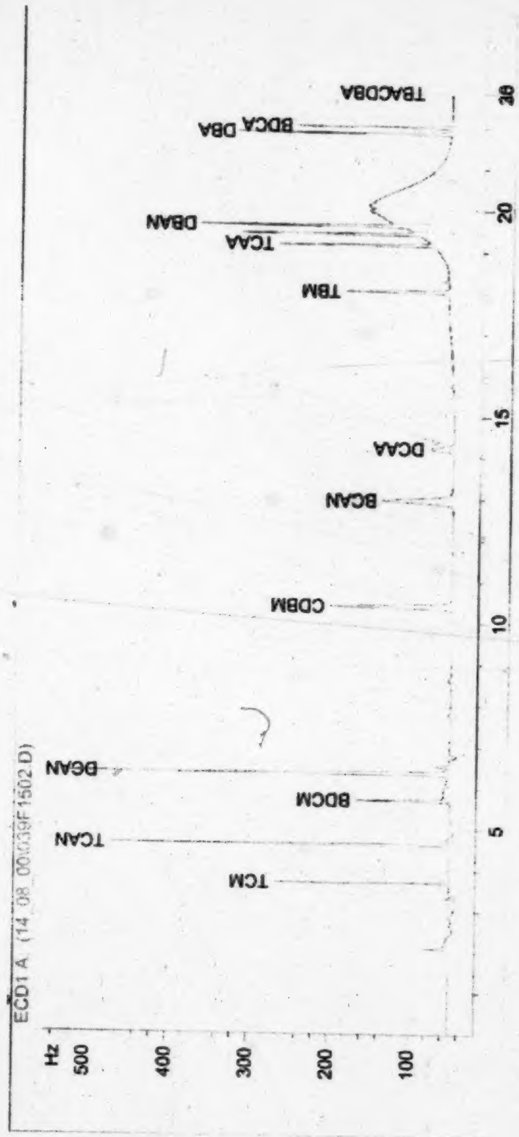


	Standard Deviation			
Concentration µg/L	**THM**	**DCAN**	**DCAA**	**TCAA**
2	0.47	0.49	0.24	0.23
5	1.22	0.97	0.30	0.29
10	1.31	1.59	0.48	0.60

MDL calculations		
Compound	**S_o (from graph)**	**MDL (µg/L)**
THM	0.4524	1.4
DCAN	0.2453	0.74
DCAA	0.1636	0.49
TCAA	0.0989	0.30

Method detection limit at 95% confidence = $3S_o$

S_o = std dev. at conc. = 0 µg/L



Sample chromatogram: 10 µg/L

Problems With And Solutions For Edmonton Daily HAA Data

Introduction

It was found by a source external to the University that several points of Edmonton daily haloacetic acid (HAA) data were unaccountably high. Therefore, a full review of all the Edmonton daily HAA data was conducted. This review included reintegration of individual peaks on chromatograms and the reexamination of calibration curves for all compounds, as well as the development of a correction factor for the HAA data. The first two items (reintegration and recalibration) yielded some change in the data, however the correction factor yielded quite significant change. This suggests that factors besides integration and calibration had a greater impact on the final analysis of the compounds.

The non-HAA data (chloroform and dichloroacetonitrile) are well within the boundaries of expected values and careful review of the non-HAA data suggests that the problem does not reside with these data. It would seem therefore, that the problem resides exclusively with the HAA data and indeed, these are the data in question. In particular, it is expected that elements of the HAA sample preparation have influenced the final chromatographic analysis. The chromatographic analysis method for both the non-HAA and the HAA samples was identical and non-HAA compound peaks were identified and quantified in the HAA sample runs. Area counts for the dichloroacetonitrile (DCAN) peaks in the HAA chromatographic runs were different from area counts for DCAN peaks in the non-HAA chromatographic runs. This suggests that some aspect of the difference between the non-HAA and HAA sample preparation methods influenced the HAA samples before they were analyzed chromatographically.

Sample Preparation

Each sample was prepared both for non-HAA and HAA analysis. Non-HAA samples underwent a solvent extraction procedure followed by direct transfer of the solvent layer accompanied by the extracted compounds into autosampler vials via Pasteur pipette. HAA samples underwent a similar solvent extraction procedure; however, the HAA extraction procedure included the addition of concentrated H_2SO_4. This was followed by transfer of the solvent layer containing the analytes of interest to a methylation vial via Pasteur pipette. Acidified methanol (10% solution of H_2SO_4 in methanol) was added and the sample was heated at 50 °C for one hour. The sample was then allowed to cool. The cooled sample was neutralized with a saturated solution of sodium bicarbonate, added dropwise. The aqueous and solvent phases were allowed to separate and the solvent layer containing the analytes of interest was transferred to an autosampler vial. The HAA sample preparation method entails several more steps than the non-HAA preparation method. In each of the steps of the HAA sample preparation method heat was generated or added (at extraction with the addition of concentrated H_2SO_4 to the water sample, during methylation with the addition of heat for one hour, and during neutralization in an exothermic, acid-base reaction). The additional steps also necessitated additional transfer procedures.

The addition and generation of heat as well as the many transfers involved in the HAA sample preparation method may have resulted in the loss of solvent due to evaporation.

The loss of solvent due to evaporation may have introduced a concentration effect in the case of relatively non-volatile compounds. To investigate this hypothesis, we must look at the area counts for the chromatographic peaks of the two compounds that were detected in both the non-HAA and HAA chromatographic runs: chloroform (TCM) and DCAN. DCAN peaks in the HAA runs have consistently higher area counts than DCAN peaks in the non-HAA runs. This is consistent with the hypothesis of a concentration effect resulting from solvent loss by evaporation due to the addition or generation of heat and multiple transfers of solvent in the HAA sample preparation method. The TCM peak area counts in the HAA runs do not show a similar or consistent trend. This can be explained by examining the relative volatilities of TCM and DCAN. The Henry's Law constants for TCM and DCAN are estimated to be 326 and 0.38 Pa m^3 mol^{-1}, respectively. The Henry's Law constant (H) is also known as the partition coefficient between the atmosphere and water and is calculated as:

H = [compound in the atmosphere at equilibrium]/[compound in water at equilibrium]

The Henry's Law constant is then a ratio of the equilibrium concentrations of the compound of interest in air and in water. The Henry's Law constant gives a measure of whether a compound is more likely to be found in air or in water at equilibrium or in other words, the volatility of the compound. The greater the Henry's Law constant for a compound, the more volatile a compound is said to be. The assumption is made here that the compounds of interest will exhibit similar trends in behaviour in the air/MTBE system found in the HAA sample preparation method as they would in an air/water system.

Since TCM is much more volatile than DCAN, it is possible that TCM would be subject to evaporation due to the same factors as the solvent loss by evaporation (addition/generation of heat, multiple transfers). If this is the case, TCM would not be subject to the concentration effect due to solvent loss by evaporation since there may be a concurrent loss of TCM due to evaporation as a result as the same processes causing the evaporation of the solvent. The lesser volatility of DCAN relative to TCM suggests that DCAN will not be lost by evaporation to the same degree during the HAA sample preparation procedure and will be subject to a concentration effect as the result of solvent loss by evaporation. The Henry's Law constants for dichloroacetic acid (DCAA) and trichloroacetic acid (TCAA) (estimated at 0.007 and 0.002 Pa m^3 mol^{-1}, respectively[1]) suggest that the HAAs will be subject to a similar concentration effect acting upon DCAN.

---

[1] Henry's Law constants for DCAA and TCAA have not been experimentally determined. The values presented here are estimates based on bond strength generated for the unionized forms of the acids.

Solution: DCAN Correction Factor

The problem of the high HAA concentration has been hypothesized to be a concentration effect resulting from solvent loss by evaporation during the HAA sample preparation procedure. To find a solution to this problem, we look to the compound DCAN and the chromatographic method used for both the non-HAA and HAA runs. As mentioned before, DCAN and TCM were detected and quantified in the HAA runs as well as in the non-HAA runs. DCAN was subject to the same concentration effect as the HAAs whereas TCM was not. Therefore, the focus must be on DCAN.

We know that each sample was prepared using both the non-HAA and the HAA methods and run chromatographically with the same method. We hypothesize that the HAA samples were subject to a concentration effect as explained previously. We can be reasonably certain that the non-HAA data are accurate by virtue of careful examination of the data, including reintegration, recalibration, comparison with historical data, and the results of a split sample analysed between two laboratories (for TCM). Therefore, it is possible to calculate a correction factor for each sample based on the area counts of DCAN from both the non-HAA and the HAA analyses. In calculating and applying this correction factor, DCAN is used as a quasi-surrogate standard in the HAA analysis to measure and correct for the increased peak areas of the HAAs due to the concentration effect explained previously. It is not strictly a true surrogate standard because the quantity of DCAN varies from sample to sample and DCAN is normally found in the samples, not added during sample preparation. Further explanation of surrogate standards is offered in the next section. In using DCAN as a surrogate standard then, a correction factor is calculated as a ratio of the DCAN area counts from the non-HAA analysis divided by the DCAN area counts from the HAA analysis. The correction factor is calculated for each sample individually, including calibration standards. Due to the multiple opportunities for solvent loss by evaporation during sample preparation and the resulting variability of the solvent loss by evaporation, it is reasonable to expect that the concentration effect will not be the same for every sample. This suggests that the use of a correction factor that can be applied to each sample individually is warranted. Once the correction factor is calculated, the measured HAA concentrations are multiplied by the calculated correction factor to give the corrected HAA values shown in the Excel spreadsheet.

The premise of this correction factor is based on the volatility of the compounds being analyzed. Of particular interest is the 1000-fold difference between the Henry's Law constants for DCAN and TCM. However, the differences between the Henry's Law constants for the acids and DCAN should also be noted. The Henry's Law constants for DCAA and TCAA are approximately 50 and 200 times lower, respectively, than that of DCAN. This suggests that the behaviour of DCAN will be more similar to the acids than to TCM, but it will not be the same. According to the Henry's Law constants, DCAN is more volatile than the acids and as a result may be subject to less of a concentration effect than will the acids. This is further evidence that DCAN is not an ideal surrogate standard as there is likely still some bias that can not be accounted for using the DCAN correction factor. In this case, the bias will be to a greater concentration of the acids than DCAN, resulting in higher concentration measurements for the acids. The magnitude of this bias cannot be determined. However, based on Henry's Law constants the bias is expected to be greater for TCAA than for DCAA. The bias in the data generated by these differences in Henry's Law constants is not expected to be so great as to render the data unusable; however it does warrant a note of consideration.

Investigating Accuracy: Comparison with Other Laboratories

Two characteristics of the data, precision and accuracy, are investigated in this and the next section. Accuracy is defined here a measure of how close the measured values are to the "true" values. In statistical terms, accuracy is sometimes known as the absence of bias. Precision is defined here as a measure of the repeatability of the data or how the data are scattered around their mean value. In statistical terms, precision is sometimes known as the variation, variability, or spread of the data.

To ensure the accuracy of the results, ideally both an internal standard and a surrogate standard would have been added to all the samples during preparation. The characteristics of an ideal internal standard include: the internal standard is not normally present in the sample, it does not cause matrix effects, and it has similar physico-chemical properties (including boiling point, retention time, detection behaviour, etc.) to the analytes of interest. To use an internal standard, a known quantity of internal standard is added to the samples after preparation and before chromatographic analysis and is quantified in the chromatographic analysis with the analytes of interest. Because the quantity of internal standard is known to be equal in all samples (if laboratory technique is of high quality) any variations in the analysed quantity of internal standard are an indication of errors in the chromatographic analysis. The internal standard also provides a method of compensating for that error by providing a known quantity of compound in each sample that can be used to correct the quantitation of the analytes of interest in each sample.

A surrogate standard works in a similar way to an internal standard; however, a known quantity of the surrogate standard is added to the pre-prepared sample and is subject to the preparation procedure as well as the chromatographic analysis. The surrogate standard can then be used to monitor for errors in the whole analysis method.

While many compounds were tested for use as an internal standard and a surrogate standard, none were found that were suitable. The primary problem for many of the compounds tested was that they coeluted with the analytes under investigation during chromatographic analysis. Coelution of an internal standard or a surrogate standard and an analyte of interest determines that neither the standard nor the compound of interest can be quantified with any degree of confidence. Therefore, the primary methods of determining the accuracy of the measured data are unavailable for use and alternate methods must be employed, namely a comparison of similar or split samples with other laboratories.

The corrected HAA data for two dates during the sampling period were compared to samples taken and analysed by Epcor Water Services at the same locations and dates. It is believed that the samples analysed at the University of Alberta and at Epcor Water Services for each date and location were taken consecutively from the same sampling taps and as such could be considered split samples. However, this is not confirmed since this sampling was done by employees of Epcor Water Services. The results are presented in Tables 1 and 2. For comparison, Table 3 presents the University of Alberta HAA data before the correction factor was applied.

Table 1 Epcor Water Services Data

Date	Location	DCAA (μg/L)	TCAA (μg/L)	THAA (μg/L)
September 20	Rossdale	9.1	9.6	19
	E.L. Smith	6.4	5.9	12
October 2	Rossdale	6.3	5.3	12
	E.L. Smith	5.8	4.8	11

Table 2 University of Alberta Corrected Data

Date	Location	DCAA (μg/L)	TCAA (μg/L)	THAA (μg/L)
September 20	Rossdale	7.4	11	18
	E.L. Smith	6.3	8.5	15
October 2	Rossdale	4.0	8.2	12
	E.L. Smith	2.3	6.2	8.5

Table 3 University of Alberta Uncorrected Data

Date	Location	DCAA (μg/L)	TCAA (μg/L)	THAA (μg/L)
September 20	Rossdale	11	17	28
	E.L. Smith	18	24	42
October 2	Rossdale	8.1	17	25
	E.L. Smith	2.7	7.2	9.9

It can be seen that the results of Tables 1 and 2 are generally in agreement, particularly when evaluating the THAA values. It is interesting to note several trends in the data that may be a result of method differences between the two laboratories. Historical monthly data provided by Epcor Water Services indicate a general trend to higher concentrations of DCAA compared with TCAA. The daily data from the University of Alberta laboratory indicate the opposite trend, with generally higher concentrations of TCAA compared with DCAA. If these trends are indeed a result of HAA method differences between the two laboratories, it is likely that there will always be an inherent margin of error between results of samples analysed at both laboratories.

A comparison of the three tables indicates that the application of the correction factor to the University of Alberta data results in data that are more in agreement with the Epcor Water Services data than are the uncorrected University of Alberta data. An exception to this statement are the data for the E.L. Smith sample from October 2, where the uncorrected DCAA and THAA University of Alberta data seem to agree with the Epcor Water Services data slightly better than do the corrected University of Alberta data. However, it should be noted from a comparison of the corrected and uncorrected University of Alberta data for this date that the correction factor produced little change in the concentration values, indicating perhaps that for this sample the concentration effect was not as much of a contributing factor in the determination of the results. The corrected data are within the range of expected HAA concentrations for Edmonton treated drinking water based on historical data provided by Epcor Water Services (not shown here), whereas the uncorrected data generally are not.

During the sampling period, a sample of Edmonton tap water was analysed at both the University of Alberta and at the City of Calgary Water Works as a comparison of the methods used in these two locations. The results of these analyses are presented in Table 4 with the uncorrected Edmonton data in parentheses. Note that there is no comparison with DCAN as the Calgary analysis method for THMs does not include haloacetonitriles and no samples were analysed for haloacetonitriles in Calgary.

Table 4. Calgary/University of Alberta Comparison Sample

Analysis Location	TCM (µg/L)	DCAA (µg/L)	TCAA (µg/L)
Calgary	16	8.0	8.7
Edmonton (UofA)	16 (16)	5.0 (3.5)	11 (9.1)

The TCM data agree for both analysis locations. This has important implications for the accuracy of the University of Alberta non-HAA data, particularly TCM, because of the analysis procedures used in each location. As previously mentioned, the University of Alberta laboratories employ a liquid-liquid extraction method followed by GC/ECD chromatographic analysis for the non-HAA compounds, a variation of the U.S. EPA method for the detection of non-HAA DBPs in drinking water (U.S. EPA method 551.1). The Calgary laboratories employ a purge-and-trap method followed by GC/ECD chromatographic analysis, a standard method for the analysis of volatiles in drinking water. The fact that there is a high level of agreement between the two very different methods improves the likelihood that both analysis methods produce accurate results for the non-HAA compounds as it is unlikely that any systematic errors inherent in the two different methods would generate the same final result.

The uncorrected and corrected University of Alberta DCAA data do not agree with the Calgary analysis, although the corrected DCAA value is closer to the value obtained in the Calgary analysis. The uncorrected University of Alberta TCAA data seem to agree more closely with the Calgary analysis than do the corrected data. The DCAA data would seem to support the use of the correction factor; however, the TCAA data do not.

Investigating Precision: Effect of the DCAN Correction Factor on the Variability of the Data

The precision or variability of the data is investigated in this section. Duplicate samples were taken every day at one of the four distribution sampling locations in Edmonton on a rotating basis. The duplicate samples were treated identically in all respects and therefore give an estimate of the precision of the whole sample handling procedure, from taking the sample through the sample preparation to chromatographic analysis. In other words, the duplicate samples can be used to determine the repeatability of the entire sample handling procedure. Further, a comparison of the precision of the corrected and pre-corrected data will illustrate the effect of the correction factor on the precision of the sample handling procedure.

The measure used here to determine the precision of the data is the standard deviation for duplicate measurements, and is calculated using the equation below:

$$s = (\Sigma d^2/2k)^{1/2}$$

Where s is the standard deviation of the duplicate measurements, d is the difference between the duplicate measurements, and k is the number of sets of duplicate measurements. The standard deviation for duplicates calculated here gives a measure of the repeatability or precision of the sample handling procedure based on the duplicate samples.

The standard deviation of the duplicate measurements for both the corrected and the precorrected data are given in Table 5.

Table 5. Standard deviations of the duplicate measurements before and after application of the correction factor (μg/L)

	TCM	TCM*	DCAN	DCAN*	DCAA	DCAA*	TCAA	TCAA*
E01	0.86	0.87	0.50	0.49	1.3	0.90	1.9	0.96
E02	0.56	0.56	0.07	0.07	1.6	0.47	1.7	0.69
E03	1.1	1.1	0.11	0.11	1.3	0.74	1.1	1.1
E04	0.65	0.65	0.10	0.10	1.0	0.51	1.2	0.32

The asterisked (*) columns contain the standard deviation for the duplicate measurements for the corrected data and E01 through E04 represent the four distribution system sampling locations in Edmonton. There were no duplicate samples taken at the water treatment plants.

It can be seen from the table that in the majority of cases the use of the correction factor improved the precision of the HAA data. This result makes sense bearing in mind the postulation that the concentration effect varies from sample to sample. The application of the correction factor will correct for the variability from sample to sample introduced by the concentration effect. The improved precision seen in Table 5 supports the use of a unique correction factor for each sample.

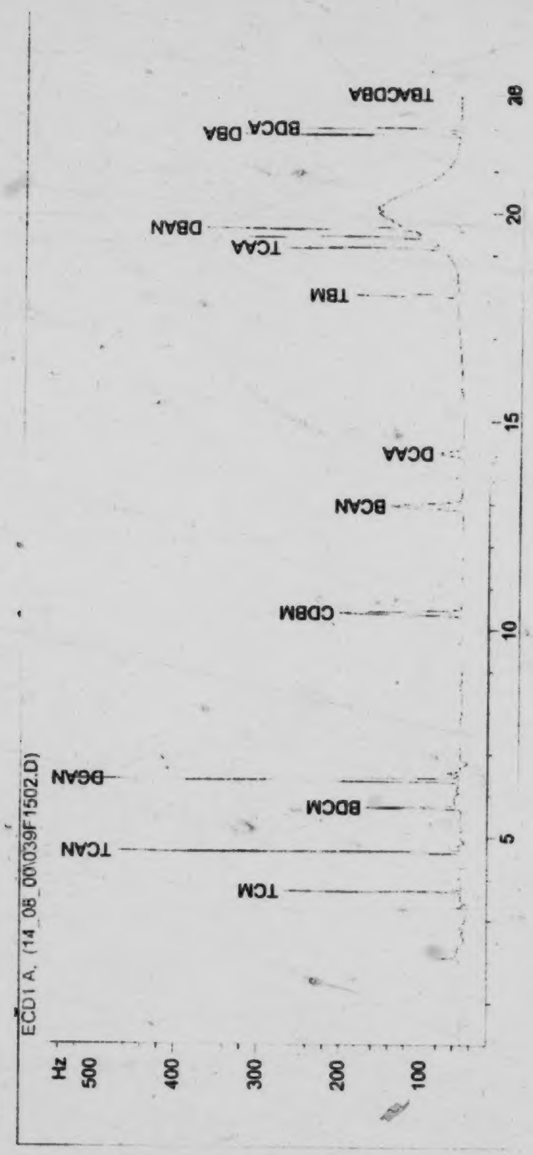
Conclusions

The hypothesis was presented that the University of Alberta daily HAA samples were subject to a concentration effect due to solvent loss by evaporation as a result of the addition and generation of heat as well as multiple transfers involved in the HAA sample preparation method. It was also postulated that this concentration effect would differ from sample to sample. To compensate for this concentration effect, a correction factor was developed using DCAN, a compound identified and quantified in both HAA and non-HAA chromatographic analyses, as a quasi-surrogate standard. This correction factor allowed for the compensation of the concentration effect on a sample-by-sample basis.

Analyses run at both Epcor Water Services and the University of Alberta on similar samples show better agreement when the correction factor is applied to the University of Alberta data than when it is not. Analysis of a split sample between the City of Calgary Water Works and the University of Alberta supports the use of the correction factor with the DCAA data. Analysis of the same split sample shows good agreement between the TCM results for the City of Calgary and the University of Alberta analysis methods.

The standard deviation of the duplicates was calculated to determine the effect of the correction factor on the precision of the measurements. In the majority of cases, the correction factor improved the precision of the data, likely by compensating for the sample-to-sample variation introduced by the concentration effect.

Based on the improvements in the accuracy and precision of the University of Alberta daily HAA data seen here, it seems reasonable to support the application of the correction factor to these data.



Sample chromatogram: 10 µg/L

Appendix for Chapters 5, 6, and 7

Tables of Correlations, Sensitivities, Specificities, and Calculated ORs from Chapters 5, 6, and 7

Chapter 5

Table A5-1. Correlations between daily DBP concentration data at water treatment plant #1 and home sampling locations in City A, and their effects on the odds ratio

TCM

	Correlation (r)	True Odds Ratio			
	TCM at WTP1 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
TCM at site A01	0.67	1.13	1.59	2.31	3.66
TCM at site A02	0.77	1.15	1.70	2.62	4.46
TCM at site A03	0.32	1.06	1.24	1.49	1.85
TCM at site A04	0.62	1.12	1.54	2.17	3.34

DCAN

	Correlation (r)	True Odds Ratio			
	TCM at WTP1 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
DCAN at site A01	0.59	1.11	1.51	2.09	3.15
DCAN at site A02	0.82	1.16	1.77	2.80	4.94
DCAN at site A03	0.46	1.09	1.37	1.77	2.42
DCAN at site A04	0.77	1.15	1.70	2.61	4.45

DCAA

	Correlation (r)	True Odds Ratio			
	TCM at WTP1 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
DCAA at site A01	0.95	1.19	1.94	3.30	6.40
DCAA at site A02	0.91	1.18	1.88	3.13	5.90
DCAA at site A03	0.86	1.17	1.82	2.94	5.33
DCAA at site A04	0.91	1.18	1.88	3.12	5.86

TCAA

	Correlation (r)	True Odds Ratio			
	TCM at WTP1 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
TCAA at site A01	0.67	1.13	1.59	2.30	3.65
TCAA at site A02	0.62	1.12	1.53	2.17	3.33
TCAA at site A03	0.50	1.10	1.42	1.88	2.67
TCAA at site A04	0.72	1.14	1.64	2.46	4.04

OR_T = the assumed OR for the true data
OR_X = the calculated OR for the measured data = OR_T^r
WTP1 = Water treatment plant #1; WTP2 = Water treatment plant #2
Sites A01-A04 are home sampling locations

Table A5-2. Correlations between **daily** DBP concentration data at water treatment plant #2 and home sampling locations in City A, and their effects on the odds ratio

TCM

True	Correlation (r)	True Odds Ratio			
	TCM at WTP2 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
TCM at site A01	0.76	1.15	1.70	2.60	4.41
TCM at site A02	0.63	1.12	1.55	2.21	3.43
TCM at site A03	0.37	1.07	1.29	1.59	2.05
TCM at site A04	0.64	1.12	1.55	2.22	3.45

DCAN

True	Correlation (r)	True Odds Ratio			
	DCAN at WTP2 (measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
DCAN at site A01	0.62	1.12	1.53	2.17	3.33
DCAN at site A02	0.76	1.15	1.69	2.59	4.39
DCAN at site A03	0.56	1.11	1.47	2.01	2.95
DCAN at site A04	0.73	1.14	1.65	2.48	4.11

DCAA

True	Correlation (r)	True Odds Ratio			
	DCAA at WTP2 (measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
DCAA at site A01	0.90	1.18	1.87	3.09	5.77
DCAA at site A02	0.88	1.17	1.84	3.02	5.55
DCAA at site A03	0.81	1.16	1.76	2.77	4.86
DCAA at site A04	0.87	1.17	1.83	2.98	5.45

TCAA

True	Correlation (r)	True Odds Ratio			
	TCAA at WTP2 (measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
TCAA at site A01	0.62	1.12	1.54	2.17	3.34
TCAA at site A02	0.64	1.12	1.56	2.24	3.49
TCAA at site A03	0.48	1.09	1.39	1.82	2.53
TCAA at site A04	0.62	1.12	1.54	2.17	3.34

OR_T = the assumed OR for the true data
OR_X = the calculated OR for the measured data = OR_T^r
WTP1 = Water treatment plant #1; WTP2 = Water treatment plant #2
Sites A01-A04 are home sampling locations

Table A5-3. Correlations between **daily** DBP concentration data at the water treatment plant and home sampling locations in City B, and their effects on the odds ratio

TCM

	Correlation (r)	True Odds Ratio			
	TCM at WTP (Measured)	$OR_T=1.20$	$OR_T=2.00$	$OR_T=3.50$	$OR_T=7.00$
		OR_X	OR_X	OR_X	OR_X
TCM at site C01	-0.09	0.98	0.94	0.90	0.84
TCM at site C02	0.11	1.02	1.08	1.14	1.23
TCM at site C03	-0.26	0.95	0.83	0.72	0.60
TCM at site C04	-0.18	0.97	0.89	0.80	0.71

BDCM

	Correlation (r)	True Odds Ratio			
	TCM at WTP (Measured)	$OR_T=1.20$	$OR_T=2.00$	$OR_T=3.50$	$OR_T=7.00$
		OR_X	OR_X	OR_X	OR_X
BDCM at site C01	-0.26	0.95	0.83	0.72	0.60
BDCM at site C02	-0.12	0.98	0.92	0.86	0.80
BDCM at site C03	-0.29	0.95	0.82	0.70	0.57
BDCM at site C04	0.12	1.02	1.08	1.16	1.25

DCAA

	Correlation (r)	True Odds Ratio			
	TCM at WTP (Measured)	$OR_T=1.20$	$OR_T=2.00$	$OR_T=3.50$	$OR_T=7.00$
		OR_X	OR_X	OR_X	OR_X
DCAA at site C01	0.03	1.01	1.02	1.04	1.06
DCAA at site C02	0.14	1.03	1.10	1.20	1.32
DCAA at site C03	0.10	1.02	1.07	1.13	1.21
DCAA at site C04	0.27	1.05	1.20	1.40	1.68

TCAA

	Correlation (r)	True Odds Ratio			
	TCM at WTP (Measured)	$OR_T=1.20$	$OR_T=2.00$	$OR_T=3.50$	$OR_T=7.00$
		OR_X	OR_X	OR_X	OR_X
TCAA at site C01	0.26	1.05	1.20	1.39	1.67
TCAA at site C02	0.66	1.13	1.58	2.29	3.61
TCAA at site C03	0.45	1.08	1.36	1.75	2.38
TCAA at site C04	0.73	1.14	1.66	2.51	4.17

OR_T = the assumed OR for the true data
OR_X = the calculated OR for the measured data = OR_T^r
WTP=Water treatment plant water treatment plant
Sites C01-C04 are home sampling locations

Table A5-4. Correlations between **monthly** DBP concentration data at water treatment plant #1 and water treatment plant #2 and distribution system locations in City A, and their effects on the odds ratio

TTHM (Water treatment plant #1)

	Correlation (r)	True Odds Ratio			
	TTHM at WTP1 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
TTHM at site 1	0.86	1.17	1.82	2.94	5.33
TTHM at site 2	0.72	1.14	1.65	2.46	4.06
TTHM at site 3	0.71	1.14	1.64	2.43	3.98
TTHM at site 4	0.80	1.16	1.74	2.72	4.74
TTHM at site 5	0.75	1.15	1.68	2.56	4.30
TTHM at site 6	0.77	1.15	1.71	2.62	4.47
TTHM at site 7	0.88	1.17	1.84	3.01	5.54
TTHM at site 8	0.79	1.15	1.73	2.69	4.65
TTHM at site 9	0.72	1.14	1.65	2.46	4.06
TTHM at site 10	0.75	1.15	1.68	2.56	4.30
TTHM at site 11	0.73	1.14	1.66	2.50	4.14
TTHM at site 12	0.89	1.18	1.85	3.05	5.65

TTHM (Water treatment plant #2)

	Correlation (r)	True Odds Ratio			
	TTHM at WTP2 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
TTHM at site 1	0.69	1.13	1.61	2.37	3.83
TTHM at site 2	0.96	1.19	1.95	3.33	6.48
TTHM at site 3	0.92	1.18	1.89	3.17	5.99
TTHM at site 4	0.96	1.19	1.95	3.33	6.48
TTHM at site 5	0.76	1.15	1.69	2.59	4.39
TTHM at site 6	0.74	1.14	1.67	2.53	4.22
TTHM at site 7	0.81	1.16	1.75	2.76	4.84
TTHM at site 8	0.97	1.19	1.96	3.37	6.60
TTHM at site 9	0.95	1.19	1.93	3.29	6.35
TTHM at site 10	0.97	1.19	1.96	3.37	6.60
TTHM at site 11	0.95	1.19	1.93	3.29	6.35
TTHM at site 12	0.80	1.16	1.74	2.72	4.74

OR_T = the assumed OR for the true data

OR_X = the calculated OR for the measured data = $OR_T{}^r$

WTP1 = Water treatment plant #1; WTP2 = Water treatment plant #2

Sites 1-12 are in the distribution system

Table A5-5. Correlations between **monthly** DBP concentration data at the water treatment plant and distribution system locations in City B, and their effects on the odds ratio

TCM

	Correlation (r)	True Odds Ratio			
	TCM at WTP (Measured)	$OR_T=1.20$	$OR_T=2.00$	$OR_T=3.50$	$OR_T=7.00$
		OR_X	OR_X	OR_X	OR_X
TCM at site 1	0.96	1.19	1.94	3.32	6.44
TCM at site 2	0.94	1.19	1.92	3.24	6.20
TCM at site 3	0.97	1.19	1.96	3.37	6.59
TCM at site 4	0.97	1.19	1.96	3.36	6.58

BDCM

	Correlation (r)	True Odds Ratio			
	BDCM at WTP (Measured)	$OR_T=1.20$	$OR_T=2.00$	$OR_T=3.50$	$OR_T=7.00$
		OR_X	OR_X	OR_X	OR_X
BDCM at site 1	0.82	1.16	1.77	2.81	4.97
BDCM at site 2	0.67	1.13	1.59	2.32	3.69
BDCM at site 3	0.82	1.16	1.76	2.79	4.91
BDCM at site 4	0.95	1.19	1.93	3.28	6.31

TTHM

	Correlation (r)	True Odds Ratio			
	TTHM at WTP (Measured)	$OR_T=1.20$	$OR_T=2.00$	$OR_T=3.50$	$OR_T=7.00$
		OR_X	OR_X	OR_X	OR_X
TTHM at site 1	0.94	1.19	1.91	3.23	6.17
TTHM at site 2	0.53	1.10	1.44	1.94	2.79
TTHM at site 3	0.56	1.11	1.47	2.01	2.97
TTHM at site 4	0.56	1.11	1.47	2.00	2.94

OR_T = the assumed OR for the true data
OR_X = the calculated OR for the measured data = OR_T'
WTP=Water treatment plant
Sites 1-4 are in the distribution system

Table A5-6, continued. Correlations between **monthly** DBP concentration data at water treatment plant and distribution system locations in City B, and their effects on the odds ratio

DCAA

	Correlation (r)	True Odds Ratio			
	DCAA at WTP (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
DCAA at site 1	0.96	1.19	1.95	3.34	6.51
DCAA at site 2	0.99	1.20	1.99	3.46	6.88
DCAA at site 3	0.98	1.20	1.98	3.43	6.79
DCAA at site 4	0.95	1.19	1.93	3.29	6.36

TCAA

	Correlation (r)	True Odds Ratio			
	TCAA at WTP (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
TCAA at site 1	0.95	1.19	1.93	3.27	6.30
TCAA at site 2	0.98	1.20	1.97	3.42	6.75
TCAA at site 3	0.97	1.19	1.95	3.35	6.55
TCAA at site 4	0.99	1.20	1.98	3.44	6.83

THAA

	Correlation (r)	True Odds Ratio			
	THAA at WTP (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
THAA at site 1	0.96	1.19	1.95	3.33	6.49
THAA at site 2	0.99	1.20	1.98	3.45	6.84
THAA at site 3	0.98	1.20	1.97	3.41	6.73
THAA at site 4	0.98	1.20	1.97	3.42	6.75

OR_T = the assumed OR for the true data
OR_X = the calculated OR for the measured data = OR_T^r
WTP=Water treatment plant
Sites 1-4 are in the distribution system

Table A5-7. Correlations between **monthly** THM concentration data at City C sampling point #2 and other distribution system locations, and their effects on the odds ratio

TTHM

	Correlation (r)	True Odds Ratio			
	TTHM at Sampling point #2 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
TTHM at site 3	-0.580	0.90	0.67	0.48	0.32
TTHM at site 4	-0.045	0.99	0.97	0.95	0.92
TTHM at site 5	-0.016	1.00	0.99	0.98	0.97
TTHM at site 6	0.099	1.02	1.07	1.13	1.21
TTHM at site 7	0.290	1.05	1.22	1.44	1.76
TTHM at site 8	0.244	1.05	1.18	1.36	1.61
TTHM at site 9	0.246	1.05	1.19	1.36	1.61
TTHM at site 10	0.374	1.07	1.30	1.60	2.07
TTHM at site 11	0.425	1.08	1.34	1.70	2.29
TTHM at site 12	0.393	1.07	1.31	1.64	2.15
TTHM at site 13	0.393	1.07	1.31	1.64	2.15
TTHM at site 14	0.354	1.07	1.28	1.56	1.99
TTHM at site 15	0.342	1.06	1.27	1.53	1.95

TCM

	Correlation (r)	True Odds Ratio			
	TCM at sampling point #2 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
TCM at site 3	-0.730	0.88	0.60	0.40	0.24
TCM at site 4	0.004	1.00	1.00	1.01	1.01
TCM at site 5	-0.031	0.99	0.98	0.96	0.94
TCM at site 6	0.182	1.03	1.13	1.26	1.42
TCM at site 7	0.344	1.06	1.27	1.54	1.95
TCM at site 8	0.283	1.05	1.22	1.43	1.73
TCM at site 9	0.374	1.07	1.30	1.60	2.07
TCM at site 10	0.414	1.08	1.33	1.68	2.24
TCM at site 11	0.461	1.09	1.38	1.78	2.45
TCM at site 12	0.423	1.08	1.34	1.70	2.28
TCM at site 13	0.419	1.08	1.34	1.69	2.26
TCM at site 14	0.383	1.07	1.30	1.62	2.11
TCM at site 15	0.452	1.09	1.37	1.76	2.41

OR_T = the assumed OR for the true data
OR_X = the calculated OR for the measured data = OR_T^r
Sites 1-15 are in the distribution system

Table A5-8, continued. Correlations between **monthly** THM concentration data at City C sampling point #2 and other distribution system locations, and their effects on the odds ratio

	BDCM				
	Correlation (r)	True Odds Ratio			
	BDCM at sampling point #2 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
BDCM at site 3	-0.490	0.91	0.71	0.54	0.39
BDCM at site 4	0.024	1.00	1.02	1.03	1.05
BDCM at site 5	-0.003	1.00	1.00	1.00	0.99
BDCM at site 6	0.209	1.04	1.16	1.30	1.50
BDCM at site 7	0.120	1.02	1.09	1.16	1.26
BDCM at site 8	0.066	1.01	1.05	1.09	1.14
BDCM at site 9	0.188	1.03	1.14	1.27	1.44
BDCM at site 10	0.304	1.06	1.23	1.46	1.81
BDCM at site 11	0.261	1.05	1.20	1.39	1.66
BDCM at site 12	0.285	1.05	1.22	1.43	1.74
BDCM at site 13	0.275	1.05	1.21	1.41	1.71
BDCM at site 14	0.239	1.04	1.18	1.35	1.59
BDCM at site 15	0.296	1.06	1.23	1.45	1.78

OR_T = the assumed OR for the true data
OR_X = the calculated OR for the measured data = OR_T'
Sites 1-15 are in the distribution system

Table A5-9. Correlations between **monthly** HAA concentration data at City C sampling point #2 and other distribution system locations, and their effects on the odds ratio

THAA

	Correlation (r)	True Odds Ratio			
	THAA at sampling point #2 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
THAA at site 3	0.01	1.00	1.01	1.01	1.02
THAA at site 4	0.79	1.15	1.73	2.69	4.65
THAA at site 5	0.61	1.12	1.53	2.15	3.28
THAA at site 6	0.61	1.12	1.53	2.15	3.28
THAA at site 7	0.54	1.10	1.45	1.97	2.86
THAA at site 8	0.54	1.10	1.45	1.97	2.86
THAA at site 9	0.62	1.12	1.54	2.17	3.34
THAA at site 10	0.58	1.11	1.49	2.07	3.09
THAA at site 11	0.79	1.15	1.73	2.69	4.65
THAA at site 12	0.71	1.14	1.64	2.43	3.98
THAA at site 13	0.71	1.14	1.64	2.43	3.98
THAA at site 14	0.81	1.16	1.75	2.76	4.84
THAA at site 15	0.58	1.11	1.49	2.07	3.09

DCAA

	Correlation (r)	True Odds Ratio			
	DCAA at sampling point #2 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
DCAA at site 3	-0.88	0.85	0.54	0.33	0.18
DCAA at site 4	-0.21	0.96	0.86	0.77	0.66
DCAA at site 5	0.39	1.07	1.31	1.63	2.14
DCAA at site 6	0.24	1.04	1.18	1.35	1.60
DCAA at site 7	0.27	1.05	1.21	1.40	1.69
DCAA at site 8	0.52	1.10	1.43	1.92	2.75
DCAA at site 9	0.62	1.12	1.54	2.17	3.34
DCAA at site 10	0.49	1.09	1.40	1.85	2.59
DCAA at site 11	0.51	1.10	1.42	1.89	2.70
DCAA at site 12	0.41	1.08	1.33	1.67	2.22
DCAA at site 13	0.43	1.08	1.35	1.71	2.31
DCAA at site 14	0.53	1.10	1.44	1.94	2.80
DCAA at site 15	0.37	1.07	1.29	1.59	2.05

OR_T = the assumed OR for the true data

OR_X = the calculated OR for the measured data = OR_T^r

Sites 1-15 are in the distribution system

Table A5-9, continued. Correlations between **monthly** HAA concentration data at City C sampling point #2 and other distribution system locations, and their effects on the odds ratio

TCAA					
	Correlation (r)	True Odds Ratio			
	TCAA at sampling point #2 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
TCAA at site 3	-0.63	0.89	0.65	0.45	0.29
TCAA at site 4	0.85	1.17	1.80	2.90	5.23
TCAA at site 5	0.50	1.10	1.41	1.87	2.65
TCAA at site 6	0.49	1.09	1.40	1.85	2.59
TCAA at site 7	0.62	1.12	1.54	2.17	3.34
TCAA at site 8	0.84	1.17	1.79	2.86	5.13
TCAA at site 9	0.60	1.12	1.52	2.12	3.21
TCAA at site 10	0.63	1.12	1.55	2.20	3.41
TCAA at site 11	0.75	1.15	1.68	2.56	4.30
TCAA at site 12	0.73	1.14	1.66	2.50	4.14
TCAA at site 13	0.63	1.12	1.55	2.20	3.41
TCAA at site 14	0.82	1.16	1.77	2.79	4.93
TCAA at site 15	0.81	1.16	1.75	2.76	4.84

OR_T = the assumed OR for the true data

OR_X = the calculated OR for the measured data = OR_T^r

Sites 1-15 are in the distribution system

Table A5-10. Correlations between **monthly** THM concentration data at City C sampling point #3 and other distribution system locations, and their effects on the odds ratio

TTHM

	Correlation (r)	True Odds Ratio			
	TTHM at sampling point #3 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
TTHM at site 4	0.78	1.15	1.72	2.67	4.60
TTHM at site 5	0.56	1.11	1.47	2.02	2.97
TTHM at site 6	0.82	1.16	1.76	2.78	4.88
TTHM at site 7	0.51	1.10	1.43	1.90	2.72
TTHM at site 8	0.52	1.10	1.43	1.92	2.75
TTHM at site 9	0.54	1.10	1.46	1.97	2.87
TTHM at site 10	0.47	1.09	1.38	1.80	2.49
TTHM at site 11	0.61	1.12	1.52	2.14	3.25
TTHM at site 12	0.74	1.14	1.67	2.53	4.23
TTHM at site 13	0.67	1.13	1.59	2.32	3.69
TTHM at site 14	0.60	1.12	1.52	2.13	3.24
TTHM at site 15	0.46	1.09	1.38	1.78	2.46

TCM

	Correlation (r)	True Odds Ratio			
	TCM at sampling point #3 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
TCM at site 4	1.00	1.20	2.00	3.50	7.00
TCM at site 5	0.69	1.13	1.61	2.36	3.79
TCM at site 6	0.80	1.16	1.74	2.72	4.72
TCM at site 7	0.79	1.15	1.73	2.69	4.65
TCM at site 8	0.62	1.12	1.53	2.16	3.31
TCM at site 9	0.67	1.13	1.59	2.32	3.69
TCM at site 10	0.63	1.12	1.55	2.21	3.43
TCM at site 11	0.47	1.09	1.38	1.79	2.47
TCM at site 12	0.69	1.13	1.61	2.36	3.80
TCM at site 13	0.57	1.11	1.49	2.04	3.04
TCM at site 14	0.45	1.09	1.37	1.76	2.40
TCM at site 15	0.34	1.06	1.27	1.53	1.94

OR_T = the assumed OR for the true data
OR_X = the calculated OR for the measured data = OR_T^r
Sites 1-15 are in the distribution system

Table A5-10, continued. Correlations between **monthly** THM concentration data at City C sampling point #3 and other distribution system locations, and their effects on the odds ratio

BDCM

	Correlation (r)	True Odds Ratio			
	BDCM at sampling point #3 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
BDCM at site 4	-1.00	0.83	0.50	0.29	0.14
BDCM at site 5	0.57	1.11	1.49	2.05	3.06
BDCM at site 6	0.68	1.13	1.60	2.34	3.76
BDCM at site 7	0.66	1.13	1.58	2.29	3.63
BDCM at site 8	0.56	1.11	1.47	2.00	2.94
BDCM at site 9	0.70	1.14	1.62	2.39	3.87
BDCM at site 10	0.65	1.12	1.56	2.24	3.51
BDCM at site 11	0.62	1.12	1.54	2.18	3.35
BDCM at site 12	0.60	1.11	1.51	2.11	3.19
BDCM at site 13	0.60	1.12	1.52	2.13	3.24
BDCM at site 14	0.55	1.11	1.46	1.99	2.92
BDCM at site 15	0.40	1.08	1.32	1.65	2.19

OR_T = the assumed OR for the true data

OR_X = the calculated OR for the measured data = OR_T^r

Sites 1-15 are in the distribution system

Table A5-11. Correlations between **monthly** HAA concentration data at City C sampling point #3 and other distribution system locations, and their effects on the odds ratio

THAA

	Correlation (r)	True Odds Ratio			
	THAA at sampling point #3 (Measured)	$OR_T=1.20$	$OR_T=2.00$	$OR_T=3.50$	$OR_T=7.00$
		OR_X	OR_X	OR_X	OR_X
THAA at site 4	0.32	1.06	1.25	1.49	1.86
THAA at site 5	0.43	1.08	1.35	1.71	2.31
THAA at site 6	0.09	1.02	1.06	1.12	1.19
THAA at site 7	0.57	1.11	1.48	2.04	3.03
THAA at site 8	0.52	1.10	1.43	1.92	2.75
THAA at site 9	0.59	1.11	1.51	2.09	3.15
THAA at site 10	0.54	1.10	1.45	1.97	2.86
THAA at site 11	-0.51	0.91	0.70	0.53	0.37
THAA at site 12	-0.01	1.00	0.99	0.99	0.98
THAA at site 13	-0.15	0.97	0.90	0.83	0.75
THAA at site 14	-0.36	0.94	0.78	0.64	0.50
THAA at site 15	0.32	1.06	1.25	1.49	1.86

DCAA

	Correlation (r)	True Odds Ratio			
	DCAA at sampling point #3 (Measured)	$OR_T=1.20$	$OR_T=2.00$	$OR_T=3.50$	$OR_T=7.00$
		OR_X	OR_X	OR_X	OR_X
DCAA at site 4	-1.00	0.83	0.50	0.29	0.14
DCAA at site 5	-0.63	0.89	0.65	0.45	0.29
DCAA at site 6	-0.33	0.94	0.80	0.66	0.53
DCAA at site 7	-0.35	0.94	0.78	0.65	0.51
DCAA at site 8	-0.20	0.96	0.87	0.78	0.68
DCAA at site 9	-0.54	0.91	0.69	0.51	0.35
DCAA at site 10	-0.33	0.94	0.80	0.66	0.53
DCAA at site 11	-0.28	0.95	0.82	0.70	0.58
DCAA at site 12	-0.25	0.96	0.84	0.73	0.61
DCAA at site 13	0.05	1.01	1.04	1.06	1.10
DCAA at site 14	-0.22	0.96	0.86	0.76	0.65
DCAA at site 15	-0.21	0.96	0.86	0.77	0.66

OR_T = the assumed OR for the true data
OR_X = the calculated OR for the measured data = OR_T^r
Sites 1-15 are in the distribution system

Table A5-11, continued. Correlations between **monthly** HAA concentration data at City C sampling point #3 and other distribution system locations, and their effects on the odds ratio

TCAA					
	Correlation (r)	True Odds Ratio			
	TCAA at sampling point #3 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
TCAA at site 4	-1.00	0.83	0.50	0.29	0.14
TCAA at site 5	0.10	1.02	1.07	1.13	1.21
TCAA at site 6	-0.27	0.95	0.83	0.71	0.59
TCAA at site 7	-0.02	1.00	0.99	0.98	0.96
TCAA at site 8	0.01	1.00	1.01	1.01	1.02
TCAA at site 9	0.38	1.07	1.30	1.61	2.09
TCAA at site 10	0.20	1.04	1.15	1.28	1.48
TCAA at site 11	-0.68	0.88	0.62	0.43	0.27
TCAA at site 12	-0.14	0.97	0.91	0.84	0.76
TCAA at site 13	-0.08	0.99	0.95	0.90	0.86
TCAA at site 14	-0.49	0.91	0.71	0.54	0.39
TCAA at site 15	-0.40	0.93	0.76	0.61	0.46

OR_T = the assumed OR for the true data
OR_X = the calculated OR for the measured data = OR_T'
Sites 1-15 are in the distribution system

Table A5-12. Sensitivity and specificity between **daily** DBP concentration data at Water treatment plant #1 and distribution system locations in City A, and their effects on the odds ratio

TCM

			True Odds Ratio			
	TCM at WTP1 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
TCM at site A01	0.71	0.90	1.11	1.53	2.19	3.34
TCM at site A02	0.71	0.90	1.11	1.53	2.19	3.34
TCM at site A03	0.29	0.75	1.01	1.03	1.06	1.09
TCM at site A04	0.71	0.90	1.11	1.53	2.19	3.34

DCAN

			True Odds Ratio			
	DCAN at WTP1 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
DCAN at site A01	0.57	0.85	1.07	1.34	1.73	2.34
DCAN at site A02	0.71	0.90	1.11	1.53	2.19	3.34
DCAN at site A03	0.27	0.75	1.00	1.01	1.03	1.05
DCAN at site A04	0.83	0.95	1.15	1.73	2.72	4.64

DCAA

			True Odds Ratio			
	DCAA at WTP1 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
DCAA at site A01	1.00	1.00	1.20	2.00	3.50	7.00
DCAA at site A02	0.71	0.90	1.11	1.53	2.19	3.34
DCAA at site A03	0.83	0.95	1.15	1.73	2.72	4.64
DCAA at site A04	0.71	0.90	1.11	1.53	2.19	3.34

TCAA

			True Odds Ratio			
	TCAA at WTP1 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
TCAA at site A01	0.83	0.95	1.15	1.73	2.72	4.64
TCAA at site A02	0.83	0.95	1.15	1.73	2.72	4.64
TCAA at site A03	0.27	0.75	1.00	1.01	1.03	1.05
TCAA at site A04	0.57	0.85	1.07	1.34	1.73	2.34

OR_T = the assumed OR for the true data
OR_X = the calculated OR for the measured data
WTP1 = Water treatment plant #1; WTP2 = Water treatment plant #2
Sites A01-A04 are home sampling locations
P_N = 0.2; P_D (OR_T=1.20)=0.23; P_D (OR_T=2.00)=0.33; P_D (OR_T=3.50)=0.0.47; P_D (OR_T=7.00)=0.64;

Table A5-13. Sensitivity and specificity between **daily** DBP concentration data at Water treatment plant #2 and distribution system locations in City A, and their effects on the odds ratio

TCM

			True Odds Ratio			
	TCM at WTP2 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
TCM at site A01	0.83	0.95	1.15	1.73	2.72	4.64
TCM at site A02	0.57	0.85	1.07	1.34	1.73	2.34
TCM at site A03	0.17	0.70	0.98	0.91	0.83	0.74
TCM at site A04	0.57	0.85	1.07	1.34	1.73	2.34

DCAN

			True Odds Ratio			
	DCAN at WTP2 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
DCAN at site A01	0.57	0.85	1.07	1.34	1.73	2.34
DCAN at site A02	0.43	0.80	1.04	1.17	1.36	1.62
DCAN at site A03	0.00	0.65	0.95	1.78	0.59	0.38
DCAN at site A04	0.71	0.90	1.07	1.53	2.19	3.34

DCAA

			True Odds Ratio			
	DCAA at WTP2 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
DCAA at site A01	0.86	1.0	1.19	1.93	3.22	5.80
DCAA at site A02	0.71	0.94	1.13	1.63	2.43	3.84
DCAA at site A03	0.71	0.94	1.14	1.65	2.47	3.92
DCAA at site A04	0.67	0.89	1.10	1.65	2.07	3.92

TCAA

			True Odds Ratio			
	TCM at WTP2 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
TCAA at site A01	0.67	0.89	1.10	1.48	2.07	3.06
TCAA at site A02	0.67	0.89	1.10	1.48	2.07	3.06
TCAA at site A03	0.29	0.78	1.01	1.05	1.11	1.18
TCAA at site A04	0.50	0.84	1.06	1.27	1.58	2.04

OR_T = the assumed OR for the true data; OR_X = the calculated OR for the measured data
WTP1 = Water treatment plant #1; WTP2 = Water treatment plant #2
Sites A01-A04 are home sampling locations
P_N = 0.2; P_D (OR_T=1.20)=0.23; P_D (OR_T=2.00)=0.33; P_D (OR_T=3.50)=0.0.47; P_D (OR_T=7.00)=0.64;

Table A5-14. Sensitivity and specificity between **daily** DBP concentration data at Water treatment plant water treatment plant and distribution system locations in City B, and their effects on the odds ratio

TCM

	TCM at WTP (Measured)		True Odds Ratio			
			OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
TCM at site C01	0.29	0.81	1.02	1.08	1.17	1.28
TCM at site C02	0.29	0.77	1.01	1.04	1.09	1.15
TCM at site C03	0.29	0.77	1.01	1.04	1.09	1.15
TCM at site C04	0.14	0.71	0.98	0.9.	0.80	0.69

BDCM

	BDCM at WTP (Measured)		True Odds Ratio			
			OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
BDCM at site C01	0.14	0.76	0.98	0.92	0.85	0.76
BDCM at site C02	0.33	0.78	1.02	1.08	1.17	1.28
BDCM at site C03	0.00	0.68	0.95	0.79	0.60	0.38
BDCM at site C04	0.43	0.81	1.04	1.18	1.39	1.67

DCAA

	DCAA at WTP (Measured)		True Odds Ratio			
			OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
DCAA at site C01	0.43	0.86	1.06	1.26	1.54	0.95
DCAA at site C02	0.43	0.77	1.04	1.20	1.41	1.72
DCAA at site C03	0.29	0.76	1.01	1.04	1.09	1.15
DCAA at site C04	0.29	0.81	1.01	1.04	1.07	1.12

TCAA

	TCAA at WTP (Measured)		True Odds Ratio			
			OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
TCAA at site C01	0.43	0.81	1.04	1.18	1.39	1.67
TCAA at site C02	0.43	0.82	1.04	1.20	1.41	1.72
TCAA at site C03	0.43	0.82	1.04	1.20	1.41	1.72
TCAA at site C04	0.57	0.86	1.08	1.35	1.77	2.42

OR_T = the assumed OR for the true data
OR_X = the calculated OR for the measured data
WTP=Water treatment plant reservoir
Sites C01-C04 are home sampling locations
P_N = 0.2; P_D (OR_T=1.20)=0.23; P_D (OR_T=2.00)=0.33; P_D (OR_T=3.50)=0.0.47; P_D (OR_T=7.00)=0.64;

Table A5-15. Sensitivity and specificity between **monthly** DBP concentration data at Water treatment plant #1 and Water treatment plant #2s and distribution system locations in City A, and their effects on the odds ratio

TTHM (Water treatment plant #1)

	TTHM at WTP1 (Measured)		True Odds Ratio			
			OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
TTHM at site 1	0.54	0.97	1.14	1.64	2.41	3.61
TTHM at site 2	0.42	0.95	1.11	1.48	2.02	2.82
TTHM at site 3	0.50	0.97	1.14	1.62	2.35	3.47
TTHM at site 4	0.50	0.94	1.11	1.50	2.08	2.97
TTHM at site 5	0.45	0.94	1.10	1.47	1.99	2.78
TTHM at site 6	0.38	0.95	1.10	1.45	1.95	2.68
TTHM at site 7	0.55	0.97	1.14	1.65	2.42	3.65
TTHM at site 8	0.45	0.97	1.13	1.59	2.27	3.29
TTHM at site 9	0.42	0.94	1.10	1.44	1.94	2.67
TTHM at site 10	0.50	0.97	1.14	1.62	2.35	3.47
TTHM at site 11	0.45	0.95	1.11	1.50	2.07	2.93
TTHM at site 12	0.55	1.00	1.18	1.82	2.79	4.36

TTHM (Water treatment plant #2)

	TTHM at WTP2 (Measured)		True Odds Ratio			
			OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
TTHM at site 1	0.55	0.92	1.10	1.47	2.03	2.90
TTHM at site 2	0.75	1.00	1.19	1.89	3.05	5.17
TTHM at site 3	0.67	0.97	1.15	1.71	3.61	4.14
TTHM at site 4	0.75	1.00	1.19	1.89	3.05	5.17
TTHM at site 5	0.55	0.88	1.08	1.38	1.82	2.50
TTHM at site 6	0.46	0.89	1.07	1.34	1.71	2.27
TTHM at site 7	0.73	0.97	1.16	1.75	2.71	4.45
TTHM at site 8	0.763	1.00	1.19	1.88	3.02	5.07
TTHM at site 9	0.67	0.94	1.13	1.61	3.26	3.65
TTHM at site 10	0.67	0.95	1.14	1.64	2.44	3.80
TTHM at site 11	0.64	0.92	1.11	1.53	2.18	3.26
TTHM at site 12	0.54	0.92	1.10	1.47	2.01	2.86

OR_T = the assumed OR for the true data
OR_X = the calculated OR for the measured data
WTP1 = Water treatment plant #1; WTP2 = Water treatment plant #2
Sites 1-12 are in the distribution system
P_N = 0.2; P_D (OR_T=1.20)=0.23; P_D (OR_T=2.00)=0.33; P_D (OR_T=3.50)=0.0.47; P_D (OR_T=7.00)=0.64;

Table A5-16. Sensitivity and specificity between **monthly** THM concentration data at Water treatment plant water treatment plant and distribution system locations in City B, and their effects on the odds ratio

TTHM

	TTHM at WTP (Measured)		True Odds Ratio			
			OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
TTHM at site 1	0.71	0.90	1.11	1.53	2.19	3.34
TTHM at site 2	0.83	0.95	1.15	1.73	2.72	4.64
TTHM at site 3	0.67	0.90	1.11	1.50	2.12	3.16
TTHM at site 4	0.83	0.95	1.15	1.73	2.72	4.64

TCM

	TCM at WTP (Measured)		True Odds Ratio			
			OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
TCM at site 1	0.75	1.00	1.13	1.60	2.37	3.77
TCM at site 2	0.50	0.92	1.10	1.44	1.94	2.71
TCM at site 3	0.50	0.92	1.10	1.44	1.94	2.71
TCM at site 4	0.50	0.92	1.10	1.44	1.94	2.71

BDCM

	BDCM at WTP (Measured)		True Odds Ratio			
			OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
BDCM at site 1	0.50	0.92	1.10	1.44	1.94	2.71
BDCM at site 2	0.75	1.00	1.19	1.89	3.05	5.17
BDCM at site 3	0.50	0.92	1.10	1.44	1.94	2.71
BDCM at site 4	0.75	1.00	1.19	1.89	3.05	5.17

OR_T = the assumed OR for the true data
OR_X = the calculated OR for the measured data
WTP = Water treatment plant water treatment plant
Sites 1-4 are in the distribution system
P_N = 0.2; P_D (OR_T=1.20)=0.23; P_D (OR_T=2.00)=0.33; P_D (OR_T=3.50)=0.0.47; P_D (OR_T=7.00)=0.64;

Table A5-17. Sensitivity and specificity between **monthly** HAA concentration data at Water treatment plant water treatment plant and distribution system locations in City B, and their effects on the odds ratio

THAA

			True Odds Ratio			
	THAA at WTP (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
THAA at site 1	0.83	1.00	1.19	1.92	3.18	5.65
THAA at site 2	0.83	1.00	1.19	1.92	3.18	5.65
THAA at site 3	0.83	1.00	1.19	1.92	3.18	5.65
THAA at site 4	0.67	0.93	1.12	1.58	2.30	3.51

DCAA

			True Odds Ratio			
	DCAA at WTP (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
DCAA at site 1	0.83	1.00	1.19	1.92	3.18	5.62
DCAA at site 2	0.83	1.00	1.19	1.92	3.18	5.62
DCAA at site 3	0.83	1.00	1.19	1.92	3.18	5.62
DCAA at site 4	0.83	1.00	1.19	1.92	3.18	5.62

TCAA

			True Odds Ratio			
	TCAA at WTP (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
TCAA at site 1	0.83	1.00	1.19	1.92	3.18	5.62
TCAA at site 2	0.83	1.00	1.19	1.92	3.18	5.62
TCAA at site 3	0.67	0.93	1.12	1.58	2.30	3.51
TCAA at site 4	0.67	0.93	1.12	1.58	2.30	3.51

OR_T = the assumed OR for the true data
OR_X = the calculated OR for the measured data
WTP = Water treatment plant water treatment plant
Sites 1-4 are in the distribution system
P_N = 0.2; P_D (OR_T=1.20)=0.23; P_D (OR_T=2.00)=0.33; P_D (OR_T=3.50)=0.0.47; P_D (OR_T=7.00)=0.64;

Table A5-18. Sensitivity and specificity between **monthly** THM concentration data at City C sampling point #2 and other distribution system locations, and their effects on the odds ratio

TTHM

			True Odds Ratio			
	TTHM at sampling point #2 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
TTHM at site 3	0.00	0.65	0.95	0.78	0.59	0.38
TTHM at site 4	0.00	0.69	1.95	0.79	0.60	0.39
TTHM at site 5	0.13	0.74	1.98	0.91	0.81	0.71
TTHM at site 6	0.00	0.64	1.95	0.78	0.59	0.37
TTHM at site 7	0.31	0.80	1.02	1.09	1.18	0.30
TTHM at site 8	0.24	0.77	1.00	1.01	1.02	1.02
TTHM at site 9	0.29	0.78	1.01	1.05	1.11	1.18
TTHM at site 10	0.41	0.82	0.98	0.95	0.93	0.92
TTHM at site 11	0.33	0.75	1.01	1.06	1.11	1.19
TTHM at site 12	0.33	0.75	1.01	1.06	1.11	1.19
TTHM at site 13	0.33	0.75	1.01	1.06	1.11	1.19
TTHM at site 14	0.33	0.75	1.01	1.06	1.11	1.19
TTHM at site 15	0.39	0.81	1.04	1.16	1.32	1.56

TCM

			True Odds Ratio			
	TCM at sampling point #2 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
TCM at site 3	0.00	0.75	0.95	0.80	0.62	0.40
TCM at site 4	0.00	0.64	0.95	0.78	0.59	0.37
TCM at site 5	0.00	0.77	0.95	0.80	0.62	0.40
TCM at site 6	0.00	0.64	0.95	0.78	0.59	0.37
TCM at site 7	0.33	0.79	1.02	1.09	1.19	1.31
TCM at site 8	0.33	0.79	1.02	1.09	1.19	1.31
TCM at site 9	0.33	0.80	1.02	1.10	1.21	1.35
TCM at site 10	0.33	0.80	1.02	1.10	1.21	1.35
TCM at site 11	0.33	0.80	1.02	1.10	1.21	1.35
TCM at site 12	0.33	0.80	1.02	1.10	1.21	1.35
TCM at site 13	0.33	0.80	1.02	1.10	1.21	1.35
TCM at site 14	0.33	0.80	1.02	1.10	1.21	1.35
TCM at site 15	0.33	0.79	1.02	1.10	1.21	1.35

OR_T = the assumed OR for the true data
OR_X = the calculated OR for the measured data
Sites 1-15 are in the distribution system
P_N = 0.2; P_D (OR_T=1.20)=0.23; P_D (OR_T=2.00)=0.33; P_D (OR_T=3.50)=0.0.47; P_D (OR_T=7.00)=0.64;
NC=not calculable

Table A5-18, continued. Sensitivity and specificity between **monthly** THM concentration data at City C sampling point #2 and other distribution system locations, and their effects on the odds ratio

BDCM

			True Odds Ratio			
	BDCM at sampling point #2 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
BDCM at site 3	NC	NC	NC	NC	NC	NC
BDCM at site 4	0.00	0.55	0.95	0.79	0.60	0.39
BDCM at site 5	0.00	0.69	0.95	0.79	0.60	0.39
BDCM at site 6	0.00	0.69	0.95	0.79	0.60	0.39
BDCM at site 7	0.00	0.69	0.95	0.79	0.60	0.39
BDCM at site 8	0.00	0.69	0.98	0.93	0.86	0.78
BDCM at site 9	0.17	0.73	0.95	0.93	0.60	0.38
BDCM at site 10	0.00	0.67	0.95	0.93	0.60	0.38
BDCM at site 11	0.00	0.67	0.95	0.93	0.60	0.38
BDCM at site 12	0.00	0.67	0.95	0.93	0.60	0.38
BDCM at site 13	0.00	0.67	0.95	0.93	0.60	0.38
BDCM at site 14	0.00	0.67	0.95	0.93	0.60	0.38
BDCM at site 15	0.00	0.64	0.95	0.78	0.59	0.37

OR_T = the assumed OR for the true data
OR_X = the calculated OR for the measured data
Sites 1-15 are in the distribution system
P_N = 0.2; P_D (OR_T=1.20)=0.23; P_D (OR_T=2.00)=0.33; P_D (OR_T=3.50)=0.0.47; P_D (OR_T=7.00)=0.64;
NC=not calculable

Table A5-19. Sensitivity and specificity between **monthly** HAA concentration data at City C sampling point #2 and other distribution system locations, and their effects on the odds ratio

THAA

			True Odds Ratio			
	THAA at sampling point #2 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
THAA at site 3	0.20	0.75	0.99	0.96	0.93	0.88
THAA at site 4	0.37	0.81	1.07	1.34	1.76	2.45
THAA at site 5	0.57	0.88	1.08	1.34	1.85	2.58
THAA at site 6	0.67	0.82	1.08	1.36	1.79	2.51
THAA at site 7	0.56	0.81	1.07	1.31	1.68	2.24
THAA at site 8	0.56	0.84	1.07	1.31	1.68	2.24
THAA at site 9	0.33	0.77	1.02	1.07	1.15	1.25
THAA at site 10	0.44	0.81	1.04	1.19	1.40	1.70
THAA at site 11	0.80	0.87	1.11	1.52	2.21	3.50
THAA at site 12	0.67	0.45	1.02	1.07	1.14	1.24
THAA at site 13	0.67	0.86	1.09	1.42	1.94	2.80
THAA at site 14	0.80	0.86	1.11	1.51	2.17	3.41
THAA at site 15	0.57	0.81	1.06	1.28	1.62	2.10

DCAA

			True Odds Ratio			
	DCAA at sampling point #2 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
DCAA at site 3	0.00	0.75	0.95	0.80	0.62	0.40
DCAA at site 4	0.00	0.75	0.95	0.80	0.62	0.40
DCAA at site 5	0.50	1.00	1.17	1.80	2.74	4.20
DCAA at site 6	0.50	0.77	1.04	1.19	1.39	1.69
DCAA at site 7	0.40	0.77	1.03	1.12	1.25	1.42
DCAA at site 8	0.40	0.79	1.03	1.14	1.29	1.50
DCAA at site 9	0.40	0.79	1.03	1.14	1.29	1.50
DCAA at site 10	0.60	0.86	1.08	1.37	1.82	2.53
DCAA at site 11	0.50	0.85	1.06	1.29	1.62	2.11
DCAA at site 12	0.40	0.77	1.03	1.12	1.25	1.42
DCAA at site 13	0.20	0.71	0.99	0.94	0.88	0.81
DCAA at site 14	0.40	0.77	1.03	1.12	1.25	1.42
DCAA at site 15	0.40	0.79	1.03	1.14	1.29	1.50

OR_T = the assumed OR for the true data
OR_X = the calculated OR for the measured data
Sites 1-15 are in the distribution system
P_N = 0.2; P_D (OR_T=1.20)=0.23; P_D (OR_T=2.00)=0.33; P_D (OR_T=3.50)=0.0.47; P_D (OR_T=7.00)=0.64;NC=not calculable

Table A5-19, continued. Sensitivity and specificity between **monthly** HAA concentration data at City C sampling point #2 and other distribution system locations, and their effects on the odds ratio

TCAA						
			True Odds Ratio			
	TCAA at sampling point #2 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
TCAA at site 3	0.00	0.60	0.94	0.77	0.58	0.36
TCAA at site 4	0.67	1.00	1.18	1.86	2.94	4.80
TCAA at site 5	0.50	0.83	1.06	1.26	1.55	1.98
TCAA at site 6	0.75	0.92	1.13	1.60	2.37	3.77
TCAA at site 7	0.75	0.86	1.10	1.47	2.08	3.15
TCAA at site 8	1.00	0.93	1.16	1.78	2.95	5.69
TCAA at site 9	0.75	0.87	1.10	1.49	2.12	3.24
TCAA at site 10	0.67	0.82	1.08	1.36	1.79	2.51
TCAA at site 11	0.75	0.87	1.10	1.49	2.12	3.24
TCAA at site 12	0.60	0.85	1.08	1.36	1.78	2.45
TCAA at site 13	0.75	0.87	1.10	1.49	2.12	3.24
TCAA at site 14	0.75	0.86	1.10	1.47	2.08	3.15
TCAA at site 15	0.60	0.79	1.06	1.27	1.59	2.09

OR_T = the assumed OR for the true data
OR_X = the calculated OR for the measured data
Sites 1-15 are in the distribution system
P_N = 0.2; P_D (OR_T=1.20)=0.23; P_D (OR_T=2.00)=0.33; P_D (OR_T=3.50)=0.0.47; P_D (OR_T=7.00)=0.64;
NC=not calculable

Table A5-20. Sensitivity and specificity between **monthly** THM concentration data at City C sampling point #3 and other distribution system locations, and their effects on the odds ratio

TTHM

			True Odds Ratio			
	TTHM at sampling point #3 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
TTHM at site 4	1.00	0.88	1.14	1.68	2.69	2.05
TTHM at site 5	0.70	0.83	1.08	1.39	1.88	2.70
TTHM at site 6	1.00	0.86	1.13	1.64	2.60	4.85
TTHM at site 7	0.63	0.80	1.07	1.31	1.67	2.23
TTHM at site 8	0.75	0.82	1.09	1.41	1.93	2.85
TTHM at site 9	0.75	0.85	1.10	1.46	2.04	3.07
TTHM at site 10	0.50	0.79	1.05	1.21	1.44	1.78
TTHM at site 11	NC	NC	NC	NC	NC	NC
TTHM at site 12	NC	NC	NC	NC	NC	NC
TTHM at site 13	NC	NC	NC	NC	NC	NC
TTHM at site 14	NC	NC	NC	NC	NC	NC
TTHM at site 15	0.56	0.82	1.06	1.29	1.62	2.12

TCM

			True Odds Ratio			
	TCM at sampling point #3 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
TCM at site 4	1.00	1.00	1.20	2.00	3.50	7.00
TCM at site 5	1.00	0.75	1.10	1.50	2.25	4.00
TCM at site 6	1.00	1.00	1.20	2.00	3.50	7.00
TCM at site 7	NC	NC	NC	NC	NC	NC
TCM at site 8	0.83	1.00	1.19	1.92	3.18	5.62
TCM at site 9	NC	NC	NC	NC	NC	NC
TCM at site 10	NC	NC	NC	NC	NC	NC
TCM at site 11	NC	NC	NC	NC	NC	NC
TCM at site 12	NC	NC	NC	NC	NC	NC
TCM at site 13	NC	NC	NC	NC	NC	NC
TCM at site 14	NC	NC	NC	NC	NC	NC
TCM at site 15	NC	NC	NC	NC	NC	NC

OR_T = the assumed OR for the true data
OR_X = the calculated OR for the measured data
Sites 1-15 are in the distribution system
P_N = 0.2; P_D (OR_T=1.20)=0.23; P_D (OR_T=2.00)=0.33; P_D (OR_T=3.50)=0.0.47; P_D (OR_T=7.00)=0.64;
NC=not calculable

Table A5-20, continued. Sensitivity and specificity between **monthly** THM concentration data at City C sampling point #3 and other distribution system locations, and their effects on the odds ratio

BDCM						
			True Odds Ratio			
	BDCM at sampling point #3 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
BDCM at site 4	NC	NC	NC	NC	NC	NC
BDCM at site 5	NC	NC	NC	NC	NC	NC
BDCM at site 6	0.50	1.00	1.17	1.80	2.74	4.20
BDCM at site 7	0.00	0.80	0.95	0.81	0.63	0.41
BDCM at site 8	1.00	1.00	0.12	1.20	3.50	7.00
BDCM at site 9	NC	NC	NC	NC	NC	NC
BDCM at site 10	0.50	1.00	1.17	1.80	2.74	4.20
BDCM at site 11	1.00	1.00	1.20	2.00	3.50	7.00
BDCM at site 12	1.00	1.00	1.20	2.00	3.50	7.00
BDCM at site 13	0.50	1.00	1.17	1.80	2.74	4.20
BDCM at site 14	1.00	1.00	1.20	2.00	3.50	7.00
BDCM at site 15	NC	NC	NC	NC	NC	NC

OR_T = the assumed OR for the true data
OR_X = the calculated OR for the measured data
Sites 1-15 are in the distribution system
P_N = 0.2; P_D (OR_T=1.20)=0.23; P_D (OR_T=2.00)=0.33; P_D (OR_T=3.50)=0.0.47; P_D (OR_T=7.00)=0.64;
NC=not calculable

Table A5-21. Sensitivity and specificity between **monthly** HAA concentration data at City C sampling point #3 and other distribution system locations, and their effects on the odds ratio

THAA

			True Odds Ratio			
	THAA at sampling point #3 (Measured)		$OR_T=1.20$	$OR_T=2.00$	$OR_T=3.50$	$OR_T=7.00$
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
THAA at site 4	0.00	0.78	0.95	0.80	0.62	0.41
THAA at site 5	0.50	0.88	1.08	1.34	1.74	2.33
THAA at site 6	0.00	0.81	0.95	0.81	0.63	0.41
THAA at site 7	0.75	0.85	1.10	1.46	2.04	3.07
THAA at site 8	0.67	0.89	1.10	1.48	2.07	3.16
THAA at site 9	0.80	0.89	1.12	1.56	2.31	3.69
THAA at site 10	0.80	0.89	1.12	1.56	2.31	3.69
THAA at site 11	0.00	0.71	0.95	0.79	0.61	0.39
THAA at site 12	0.00	0.67	0.95	0.79	0.60	0.38
THAA at site 13	0.00	0.71	0.95	0.79	0.61	0.39
THAA at site 14	0.50	0.83	1.06	1.26	1.55	1.98
THAA at site 15	0.40	0.78	1.03	1.13	1.27	1.46

DCAA

			True Odds Ratio			
	DCAA at sampling point #3 (Measured)		$OR_T=1.20$	$OR_T=2.00$	$OR_T=3.50$	$OR_T=7.00$
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
DCAA at site 4	NC	NC	NC	NC	NC	NC
DCAA at site 5	0.00	0.67	0.95	0.79	0.59	0.38
DCAA at site 6	0.00	0.80	0.95	0.81	0.63	0.41
DCAA at site 7	0.00	0.80	0.95	0.81	0.63	0.41
DCAA at site 8	0.00	0.80	0.95	0.81	0.63	0.41
DCAA at site 9	0.00	0.80	0.95	0.81	0.63	0.41
DCAA at site 10	0.00	0.75	0.95	0.80	0.62	0.40
DCAA at site 11	0.00	0.75	0.95	0.80	0.62	0.40
DCAA at site 12	0.00	0.80	0.95	0.81	0.63	0.41
DCAA at site 13	0.00	0.80	0.95	0.81	0.63	0.41
DCAA at site 14	0.00	0.80	0.95	0.81	0.63	0.41
DCAA at site 15	0.00	0.80	0.95	0.81	0.63	0.41

OR_T = the assumed OR for the true data
OR_X = the calculated OR for the measured data
Sites 1-15 are in the distribution system
$P_N = 0.2$; P_D (OR_T=1.20)=0.23; P_D (OR_T=2.00)=0.33; P_D (OR_T=3.50)=0.0.47; P_D (OR_T=7.00)=0.64;
NC=not calculable

Table A5-21, continued. Sensitivity and specificity between **monthly** HAA concentration data at City C sampling point #3 and other distribution system locations, and their effects on the odds ratio

TCAA

	TCAA at sampling point #3 (Measured)		True Odds Ratio			
			OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
TCAA at site 4	NC	NC	NC	NC	NC	NC
TCAA at site 5	0.00	0.75	0.95	0.80	0.62	0.40
TCAA at site 6	0.33	0.75	1.01	1.06	1.11	1.19
TCAA at site 7	NC	NC	NC	NC	NC	NC
TCAA at site 8	0.33	0.75	1.01	1.06	1.11	1.19
TCAA at site 9	NC	NC	NC	NC	NC	NC
TCAA at site 10	1.00	0.83	1.12	1.60	2.49	4.57
TCAA at site 11	0.50	0.80	1.05	1.22	1.47	1.83
TCAA at site 12	0.50	0.80	1.05	1.22	1.47	1.83
TCAA at site 13	0.50	0.80	1.05	1.22	1.47	1.83
TCAA at site 14	0.50	0.80	1.05	1.22	1.47	1.83
TCAA at site 15	0.50	0.80	1.05	1.22	1.47	1.83

OR_T = the assumed OR for the true data
OR_X = the calculated OR for the measured data
Sites 1-15 are in the distribution system
P_N = 0.2; P_D (OR_T=1.20)=0.23; P_D (OR_T=2.00)=0.33; P_D (OR_T=3.50)=0.0.47; P_D (OR_T=7.00)=0.64;
NC=not calculable

Chapter 6

Table A6-1. Correlations between TCM and other DBP species at Water treatment plant #1 and Water treatment plant #2 water treatment plants and distribution system locations, and their effects on the odds ratio

City A Daily

	Correlation (r)	True Odds Ratio			
	TCM at WTP1 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
DCAN at WTP1	0.92	1.18	1.89	3.17	6.01
DCAA at WTP1	0.79	1.16	1.73	2.70	4.69
TCAA at WTP1	0.89	1.18	1.85	3.04	5.62
	Correlation (r)	True Odds Ratio			
	TCM at WTP2 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
DCAN at WTP2	0.86	1.17	1.81	2.93	5.31
DCAA at WTP2	0.55	1.11	1.47	2.00	2.94
TCAA at WTP2	0.66	1.13	1.58	2.30	3.64
	Correlation (r)	True Odds Ratio			
	TCM at A01 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
DCAN at site A01	0.84	1.17	1.79	2.86	5.11
DCAA at site A01	0.65	1.13	1.57	2.26	3.56
TCAA at site A01	0.83	1.16	1.77	2.82	5.00

OR_T = the assumed OR for the true data

OR_X = the calculated OR for the measured data = OR_T^r

WTP1 = Water treatment plant #1; WTP2 = Water treatment plant #2

Sites A01-A04 are home sampling locations

Table A6-2, continued. Correlations between TCM and other DBP species at Water treatment plant #1 and Water treatment plant #2 water treatment plants and distribution system locations, and their effects on the odds ratio

City A Daily

	Correlation (r)	True Odds Ratio			
	TCM at A02 (Measured)	$OR_T=1.20$	$OR_T=2.00$	$OR_T=3.50$	$OR_T=7.00$
		OR_X	OR_X	OR_X	OR_X
DCAN at site A02	0.84	1.16	1.79	2.85	5.10
DCAA at site A02	0.62	1.12	1.54	2.18	3.35
TCAA at site A02	0.75	1.15	1.69	2.57	4.33

	Correlation (r)	True Odds Ratio			
	TCM at A03 (Measured)	$OR_T=1.20$	$OR_T=2.00$	$OR_T=3.50$	$OR_T=7.00$
		OR_X	OR_X	OR_X	OR_X
DCAN at site A03	0.80	1.16	1.75	2.74	4.78
DCAA at site A03	0.62	1.12	1.54	2.18	3.35
TCAA at site A03	0.69	1.13	1.62	2.38	3.84

	Correlation (r)	True Odds Ratio			
	TCM at A04 (Measured)	$OR_T=1.20$	$OR_T=2.00$	$OR_T=3.50$	$OR_T=7.00$
		OR_X	OR_X	OR_X	OR_X
DCAN at site A04	0.81	1.16	1.75	2.76	4.84
DCAA at site A04	0.68	1.13	1.60	2.34	3.73
TCAA at site A04	0.80	1.16	1.74	2.73	4.75

OR_T = the assumed OR for the true data

OR_X = the calculated OR for the measured data = OR_T^r

WTP1 = Water treatment plant #1; WTP2 = Water treatment plant #2

Sites A01-A04 are home sampling locations

Table A6-3. Correlations between TCM and other DBP species at Water treatment plant water treatment plant and distribution system locations, and their effects on the odds ratio

City B Daily

	Correlation (r)	True Odds Ratio			
	TCM at WTP (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
BDCM at WTP	0.74	1.14	1.67	2.53	4.22
DCAA at WTP	0.41	1.08	1.33	1.67	2.22
TCAA at WTP	0.22	1.04	1.16	1.31	1.52

	Correlation (r)	True Odds Ratio			
	TCM at B01 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
BDCM at site B01	0.82	1.16	1.76	2.79	4.91
DCAA at site B01	0.72	1.14	1.65	2.46	4.05
TCAA at site B01	0.69	1.13	1.61	2.38	3.84

	Correlation (r)	True Odds Ratio			
	TCM at B02 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
BDCM at site B02	0.62	1.12	1.54	2.17	3.34
DCAA at site B02	0.42	1.08	1.34	1.69	2.26
TCAA at site B02	0.20	1.04	1.15	1.29	1.48

	Correlation (r)	True Odds Ratio			
	TCM at B03 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
BDCM at site B03	0.91	1.18	1.88	3.13	5.89
DCAA at site B03	0.17	1.03	1.12	1.23	1.38
TCAA at site B03	0.71	1.14	1.63	2.43	3.97

	Correlation (r)	True Odds Ratio			
	TCM at B04 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
BDCM at site B04	0.59	1.11	1.50	2.08	3.12
DCAA at site B04	0.01	1.00	1.01	1.01	1.02
TCAA at site B04	-0.08	0.99	0.95	0.90	0.86

OR_T = the assumed OR for the true data

OR_X = the calculated OR for the measured data = OR_T'

WTP=Water treatment plant; Sites B01-B04 are home sampling locations

Table A6-4. Correlations between TCM or TTHM and other DBP species at Water treatment plant #1 and Water treatment plant #2 water treatment plants and distribution system locations, and their effects on the odds ratio

City A Monthly

	Correlation (r)	True Odds Ratio			
	TCM at WTP1 (Measured)	$OR_T=1.20$	$OR_T=2.00$	$OR_T=3.50$	$OR_T=7.00$
		OR_X	OR_X	OR_X	OR_X
BDCM at WTP1	0.86	1.17	1.82	2.94	5.33
TTHM at WTP1	0.99	1.20	1.99	3.46	6.87
DCAA at WTP1	0.82	1.16	1.77	2.79	4.93
TCAA at WTP1	0.87	1.17	1.83	2.97	5.44
THAA at WTP1	0.86	1.17	1.82	2.94	5.33

	Correlation (r)	True Odds Ratio			
	TCM at WTP2 (Measured)	$OR_T=1.20$	$OR_T=2.00$	$OR_T=3.50$	$OR_T=7.00$
		OR_X	OR_X	OR_X	OR_X
BDCM at WTP2	0.90	1.18	1.87	3.09	5.76
TTHM at WTP2	0.97	1.19	1.96	3.37	6.60
DCAA at WTP2	0.85	1.17	1.80	2.90	5.23
TCAA at WTP2	0.73	1.14	1.66	2.50	4.14
THAA at WTP2	0.82	1.16	1.77	2.79	4.93

	Correlation (r)	True Odds Ratio			
	TTHM at WTP1 (Measured)	$OR_T=1.20$	$OR_T=2.00$	$OR_T=3.50$	$OR_T=7.00$
		OR_X	OR_X	OR_X	OR_X
BDCM at WTP1	0.83	1.16	1.78	2.83	5.03
TCM at WTP1	0.99	1.20	1.99	3.46	6.87
DCAA at WTP1	0.80	1.16	1.74	2.72	4.74
TCAA at WTP1	0.77	1.15	1.71	2.62	4.47
THAA at WTP1	0.81	1.16	1.75	2.76	4.84

	Correlation (r)	True Odds Ratio			
	TTHM at WTP2 (Measured)	$OR_T=1.20$	$OR_T=2.00$	$OR_T=3.50$	$OR_T=7.00$
		OR_X	OR_X	OR_X	OR_X
BDCM at WTP2	0.83	1.16	1.78	2.83	5.03
TCM at WTP2	0.97	1.19	1.96	3.37	6.60
DCAA at WTP2	0.75	1.15	1.68	2.56	4.30
TCAA at WTP2	0.59	1.11	1.51	2.09	3.15
THAA at WTP2	0.71	1.14	1.64	2.43	3.98

OR_T = the assumed OR for the true data

OR_X = the calculated OR for the measured data = OR_T^r

WTP1 = Water treatment plant #1; WTP2 = Water treatment plant #2

Table A6-5. Correlations between TCM and other DBP species at Water treatment plant water treatment plant and distribution system locations, and their effects on the odds ratio

City B Monthly

	Correlation (r)	True Odds Ratio			
	TCM at WTP (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
BDCM at WTP	0.86	1.17	1.81	2.92	5.29
TTHM at WTP	1.00	1.20	1.99	3.48	6.95
DCAA at WTP	0.96	1.19	1.94	3.32	6.45
TCAA at WTP	0.95	1.19	1.93	3.28	6.31
THAA at WTP	0.93	1.19	1.91	3.22	6.16

	Correlation (r)	True Odds Ratio			
	TCM at B01 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
BDCM at site B01	0.58	1.11	1.49	2.07	3.09
TTHM at site B01	1.00	1.20	2.00	3.50	7.00
DCAA at site B01	0.93	1.18	1.90	3.19	6.05
TCAA at site B01	0.93	1.18	1.90	3.19	6.06
THAA at site B01	0.94	1.19	1.92	3.26	6.27

	Correlation (r)	True Odds Ratio			
	TCM at B02 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
BDCM at site B02	0.66	1.13	1.58	2.30	3.64
TTHM at site B02	1.00	1.20	2.00	3.50	7.00
DCAA at site B02	0.95	1.19	1.93	3.28	6.33
TCAA at site B02	0.81	1.16	1.75	2.74	4.79
THAA at site B02	0.94	1.19	1.92	3.26	6.27

OR_T = the assumed OR for the true data

OR_X = the calculated OR for the measured data = OR_T^r

WTP=Water treatment plant

Sites B01-B04 are in the distribution system

Table A6-6, continued. Correlations between TCM and other DBP species at Water treatment plant water treatment plant and distribution system locations, and their effects on the odds ratio

City B Monthly

	Correlation (r)	True Odds Ratio			
	TCM at B03 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
BDCM at site B03	0.65	1.13	1.57	2.26	3.54
TTHM at site B03	1.00	1.20	2.00	3.50	7.00
DCAA at site B03	0.87	1.17	1.83	2.98	5.46
TCAA at site B03	0.93	1.19	1.91	3.22	6.14
THAA at site B03	0.93	1.18	1.90	3.19	6.06

	Correlation (r)	True Odds Ratio			
	TCM at B04 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
BDCM at site B04	0.63	1.12	1.55	2.21	3.43
TTHM at site B04	1.00	1.20	2.00	3.50	7.00
DCAA at site B04	0.91	1.18	1.87	3.11	5.83
TCAA at site B04	0.95	1.19	1.93	3.28	6.34
THAA at site B04	0.96	1.19	1.94	3.31	6.41

OR_T = the assumed OR for the true data

OR_X = the calculated OR for the measured data = OR_T^r

WTP=Water treatment plant

Sites B01-B04 are in the distribution system

Table A6-7. Correlations between TTHM and other DBP species at Water treatment plant water treatment plant and distribution system locations, and its effects on the odds ratio

City B Monthly

	Correlation (r)	True Odds Ratio			
	TTHM at WTP (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
BDCM at WTP	0.85	1.17	1.80	2.88	5.18
TCM at WTP	1.00	1.20	2.00	3.49	6.97
DCAA at WTP	0.96	1.19	1.94	3.32	6.45
TCAA at WTP	0.95	1.19	1.93	3.27	6.30
THAA at WTP	0.93	1.18	1.90	3.19	6.06

	Correlation (r)	True Odds Ratio			
	TTHM at B01 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
BDCM at B01	0.60	1.12	1.52	2.13	3.23
TCM at B01	1.00	1.20	2.00	3.50	7.00
DCAA at B01	0.94	1.19	1.92	3.24	6.22
TCAA at B01	0.94	1.19	1.92	3.25	6.25
THAA at B01	0.89	1.18	1.85	3.04	5.63

	Correlation (r)	True Odds Ratio			
	TTHM at B02 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
BDCM at B02	0.67	1.13	1.59	2.32	3.70
TCM at B02	1.00	1.20	2.00	3.50	7.00
DCAA at B02	0.60	1.11	1.51	2.11	3.18
TCAA at B02	0.49	1.09	1.40	1.85	2.59
THAA at B02	0.60	1.12	1.51	2.12	3.21

OR_T = the assumed OR for the true data
OR_X = the calculated OR for the measured data = OR_T^r
WTP=Water treatment plant
Sites B01-B04 are in the distribution system

Table A6-7, continued. Correlations between TTHM and other DBP species at Water treatment plant water treatment plant and distribution system locations, and their effects on the odds ratio

City B Monthly

	Correlation (r)	True Odds Ratio			
	TTHM at B03 (Measured)	$OR_T=1.20$	$OR_T=2.00$	$OR_T=3.50$	$OR_T=7.00$
		OR_X	OR_X	OR_X	OR_X
BDCM at site B03	0.66	1.13	1.58	2.28	3.59
TCM at site B03	1.00	1.20	2.00	3.50	7.00
DCAA at site B03	0.57	1.11	1.49	2.05	3.06
TCAA at site B03	0.70	1.14	1.62	2.40	3.90
THAA at site B03	0.64	1.12	1.55	2.22	3.45

	Correlation (r)	True Odds Ratio			
	TTHM at B04 (Measured)	$OR_T=1.20$	$OR_T=2.00$	$OR_T=3.50$	$OR_T=7.00$
		OR_X	OR_X	OR_X	OR_X
BDCM at site B04	0.64	1.12	1.56	2.24	3.49
TCM at site B04	1.00	1.20	2.00	3.50	7.00
DCAA at site B04	0.64	1.12	1.56	2.23	3.48
TCAA at site B04	0.65	1.13	1.57	2.26	3.55
THAA at site B04	0.67	1.13	1.60	2.33	3.71

OR_T = the assumed OR for the true data
OR_X = the calculated OR for the measured data = $OR_T^{\,r}$
WTP=Water treatment plant
Sites B01-B04 are in the distribution system

Table A6-8. Correlations between TTHM and other DBP species at reference points #2 and #3 and other distribution system locations, and their effects on the odds ratio

City C Monthly

	Correlation (r)	True Odds Ratio			
	TTHM at sampling point #2 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
TCM at site 2	1.00	1.20	2.00	3.50	7.00
BDCM at site 2	0.97	1.19	1.96	3.36	6.58
THAA at site 2	0.71	1.14	1.64	2.45	4.01
DCAA at site 2	0.89	1.18	1.85	3.05	5.65
TCAA at site 2	0.70	1.14	1.62	2.40	3.90

	Correlation (r)	True Odds Ratio			
	TTHM at sampling point #3 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
TCM at site 3	1.00	1.20	2.00	3.50	7.00
BDCM at site 3	0.99	1.20	1.98	3.45	6.84
THAA at site 3	0.31	1.06	1.24	1.47	1.82
DCAA at site 3	0.52	1.10	1.43	1.92	2.75
TCAA at site 3	0.36	1.07	1.28	1.57	2.01

	Correlation (r)	True Odds Ratio			
	TTHM at sampling point #4 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
TCM at site 4	1.00	1.20	2.00	3.50	7.00
BDCM at site 4	0.99	1.20	1.98	3.43	6.80
THAA at site 4	0.33	1.06	1.26	1.51	1.90
DCAA at site 4	0.58	1.11	1.49	2.07	3.09
TCAA at site 4	-0.15	0.97	0.90	0.83	0.75

OR_T = the assumed OR for the true data
OR_X = the calculated OR for the measured data = OR_T^r
Sites 1-15 are in the distribution system

Table A6-8, continued. Correlations between TTHM and other DBP species at reference points #2 and #3 and other distribution system locations, and their effects on the odds ratio

City C Monthly

	Correlation (r)	True Odds Ratio			
	TTHM at sampling point #5 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
TCM at site 5	1.00	1.20	2.00	3.50	7.00
BDCM at site 5	0.97	1.19	1.96	3.38	6.62
THAA at site 5	0.14	1.03	1.10	1.19	1.31
DCAA at site 5	-0.02	1.00	0.99	0.98	0.96
TCAA at site 5	-0.01	1.00	0.99	0.99	0.98

	Correlation (r)	True Odds Ratio			
	TTHM at sampling point #6 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
TCM at site 6	1.00	1.20	2.00	3.50	7.00
BDCM at site 6	0.97	1.19	1.96	3.37	6.59
THAA at site 6	0.26	1.05	1.20	1.39	1.67
DCAA at site 6	0.16	1.03	1.12	1.22	1.37
TCAA at site 6	0.16	1.03	1.12	1.22	1.37

	Correlation (r)	True Odds Ratio			
	TTHM at sampling point #7 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
TCM at site 7	1.00	1.20	2.00	3.50	6.99
BDCM at site 7	0.91	1.18	1.87	3.11	5.83
THAA at site 7	0.15	1.03	1.11	1.21	1.35
DCAA at site 7	0.11	1.02	1.08	1.15	1.24
TCAA at site 7	-0.06	0.99	0.96	0.93	0.89

OR_T = the assumed OR for the true data
OR_X = the calculated OR for the measured data = OR_T^r
Sites 1-15 are in the distribution system

Table A6-8, continued. Correlations between TTHM and other DBP species at reference points #2 and #3 and other distribution system locations, and their effects on the odds ratio

City C Monthly

	Correlation (r)	True Odds Ratio			
	TTHM at sampling point #8 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
TCM at site 8	1.00	1.20	2.00	3.50	6.99
BDCM at site 8	0.92	1.18	1.89	3.15	5.93
THAA at site 8	0.25	1.05	1.19	1.36	1.62
DCAA at site 8	0.26	1.05	1.20	1.39	1.66
TCAA at site 8	0.07	1.01	1.05	1.09	1.15

	Correlation (r)	True Odds Ratio			
	TTHM at sampling point #9 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
TCM at site 9	1.00	1.20	1.99	3.48	6.95
BDCM at site 9	0.85	1.17	1.80	2.89	5.20
THAA at site 9	-0.02	1.00	0.99	0.98	0.97
DCAA at site 9	0.22	1.04	1.16	1.32	1.53
TCAA at site 9	0.20	1.04	1.15	1.28	1.48

	Correlation (r)	True Odds Ratio			
	TTHM at sampling point #10 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
TCM at site 10	1.00	1.20	2.00	3.50	6.99
BDCM at site 10	0.91	1.18	1.88	3.12	5.85
THAA at site 10	0.22	1.04	1.17	1.32	1.55
DCAA at site 10	0.27	1.05	1.21	1.40	1.69
TCAA at site 10	0.04	1.01	1.03	1.05	1.08

OR_T = the assumed OR for the true data
OR_X = the calculated OR for the measured data = OR_T^r
Sites 1-15 are in the distribution system

Table A6-8, continued. Correlations between TTHM and other DBP species at reference points #2 and #3 and other distribution system locations, and their effects on the odds ratio

City C Monthly

	Correlation (r)	True Odds Ratio			
	TTHM at sampling point #11 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
TCM at site 11	1.00	1.20	2.00	3.49	6.97
BDCM at site 11	0.84	1.17	1.80	2.88	5.17
THAA at site 11	0.33	1.06	1.25	1.51	1.89
DCAA at site 11	0.29	1.05	1.22	1.44	1.76
TCAA at site 11	0.06	1.01	1.04	1.08	1.12

	Correlation (r)	True Odds Ratio			
	TTHM at sampling point #12 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
BDCM at site 12	1.00	1.20	2.00	3.50	6.99
TCM at site 12	0.89	1.18	1.86	3.06	5.70
THAA at site 12	0.42	1.08	1.34	1.70	2.27
DCAA at site 12	0.35	1.07	1.27	1.55	1.98
TCAA at site 12	0.20	1.04	1.15	1.28	1.48

	Correlation (r)	True Odds Ratio			
	TTHM at sampling point #13 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
TCM at site 13	1.00	1.20	2.00	3.50	6.99
BDCM at site 13	0.90	1.18	1.86	3.07	5.72
THAA at site 13	0.21	1.04	1.16	1.31	1.51
DCAA at site 13	0.28	1.05	1.21	1.42	1.72
TCAA at site 13	-0.02	1.00	0.99	0.98	0.96

OR_T = the assumed OR for the true data

OR_X = the calculated OR for the measured data = OR_T^r

Sites 1-15 are in the distribution system

Table A6-8, continued. Correlations between TTHM and other DBP species at reference points #2 and #3 and other distribution system locations, and their effects on the odds ratio

City C Monthly

	Correlation (r)	True Odds Ratio			
	TTHM at sampling point #14 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
TCM at site 14	1.00	1.20	2.00	3.50	6.99
BDCM at site 14	0.91	1.18	1.87	3.11	5.83
THAA at site 14	0.25	1.05	1.19	1.37	1.63
DCAA at site 14	0.26	1.05	1.20	1.39	1.66
TCAA at site 14	-0.04	0.99	0.97	0.95	0.93

	Correlation (r)	True Odds Ratio			
	TTHM at sampling point #15 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
TCM at site 15	1.00	1.20	2.00	3.50	6.99
BDCM at site 15	0.91	1.18	1.88	3.13	5.89
THAA at site 15	0.16	1.03	1.12	1.22	1.36
DCAA at site 15	0.20	1.04	1.15	1.28	1.48
TCAA at site 15	0.13	1.02	1.09	1.18	1.29

OR_T = the assumed OR for the true data

OR_X = the calculated OR for the measured data = OR_T^r

Sites 1-15 are in the distribution system

Table A6-9. Correlations between TCM and other DBP species at sampling points #2 and #3 and other distribution system locations, and their effects on the odds ratio

City C Monthly

	Correlation (r)	True Odds Ratio			
	TCM at sampling point #2 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
TTHM at site 2	1.00	1.20	2.00	3.50	7.00
BDCM at site 2	0.96	1.19	1.95	3.34	6.51
THAA at site 2	0.81	1.16	1.75	2.76	4.84
DCAA at site 2	0.89	1.18	1.85	3.05	5.65
TCAA at site 2	0.71	1.14	1.64	2.43	3.98

	Correlation (r)	True Odds Ratio			
	TCM at sampling point #3 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
TTHM at site 3	1.00	1.20	2.00	3.50	7.00
BDCM at site 3	0.99	1.20	1.98	3.44	6.81
THAA at site 3	0.51	1.10	1.42	1.89	2.70
DCAA at site 3	0.52	1.10	1.43	1.92	2.75
TCAA at site 3	0.36	1.07	1.28	1.57	2.01

	Correlation (r)	True Odds Ratio			
	TCM at sampling point #4 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
TTHM at site 4	1.00	1.20	2.00	3.50	7.00
BDCM at site 4	0.98	1.20	1.97	3.42	6.75
THAA at site 4	0.10	1.02	1.07	1.13	1.21
DCAA at site 4	0.57	1.11	1.48	2.04	3.03
TCAA at site 4	-0.14	0.97	0.91	0.84	0.76

OR_T = the assumed OR for the true data

OR_X = the calculated OR for the measured data = OR_T^r

Sites 1-15 are in the distribution system

Table A6-9, continued. Correlations between TCM and other DBP species at sampling points #2 and #3 and other distribution system locations, and their effects on the odds ratio

City C Monthly

	Correlation (r)	True Odds Ratio			
	TCM at sampling point #5 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
TTHM at site 5	1.00	1.20	2.00	3.50	7.00
BDCM at site 5	0.96	1.19	1.95	3.35	6.53
THAA at site 5	0.07	1.01	1.05	1.09	1.15
DCAA at site 5	-0.01	1.00	0.99	0.99	0.98
TCAA at site 5	0.01	1.00	1.01	1.01	1.02

	Correlation (r)	True Odds Ratio			
	TCM at sampling point #6 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
TTHM at site 6	1.00	1.20	2.00	3.50	7.00
BDCM at site 6	0.96	1.19	1.95	3.34	6.50
THAA at site 6	0.23	1.04	1.17	1.33	1.56
DCAA at site 6	0.17	1.03	1.13	1.24	1.39
TCAA at site 6	0.17	1.03	1.13	1.24	1.39

	Correlation (r)	True Odds Ratio			
	TCM at sampling point #7 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
TTHM at site 7	1.00	1.20	2.00	3.50	6.99
BDCM at site 7	0.89	1.18	1.85	3.04	5.62
THAA at site 7	0.07	1.01	1.05	1.09	1.15
DCAA at site 7	0.12	1.02	1.09	1.16	1.26
TCAA at site 7	-0.05	0.99	0.97	0.94	0.91

OR_T = the assumed OR for the true data

OR_X = the calculated OR for the measured data = OR_T^r

Sites 1-15 are in the distribution system

Table A6-9, continued. Correlations between TCM and other DBP species at sampling points #2 and #3 and other distribution system locations, and their effects on the odds ratio

City C Monthly

	Correlation (r)	True Odds Ratio			
	TCM at sampling point #8 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
TTHM at site 8	1.00	1.20	2.00	3.50	6.99
BDCM at site 8	0.90	1.18	1.86	3.08	5.73
THAA at site 8	0.23	1.04	1.17	1.33	1.56
DCAA at site 8	0.28	1.05	1.21	1.42	1.72
TCAA at site 8	0.10	1.02	1.07	1.13	1.21

	Correlation (r)	True Odds Ratio			
	TCM at sampling point #9 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
TTHM at site 9	1.00	1.20	1.99	3.48	6.95
BDCM at site 9	0.80	1.16	1.74	2.72	4.73
THAA at site 9	0.31	1.06	1.24	1.47	1.83
DCAA at site 9	0.22	1.04	1.16	1.32	1.53
TCAA at site 9	0.24	1.04	1.18	1.35	1.60

	Correlation (r)	True Odds Ratio			
	TCM at sampling point #10 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
TTHM at site 10	1.00	1.20	2.00	3.50	6.99
BDCM at site 10	0.89	1.18	1.85	3.05	5.64
THAA at site 10	0.27	1.05	1.21	1.40	1.69
DCAA at site 10	0.29	1.05	1.22	1.44	1.76
TCAA at site 10	0.07	1.01	1.05	1.09	1.15

OR_T = the assumed OR for the true data
OR_X = the calculated OR for the measured data = OR_T^r
Sites 1-15 are in the distribution system

Table A6-9, continued. Correlations between TCM and other DBP species at sampling points #2 and #3 and other distribution system locations, and their effects on the odds ratio

City C Monthly

	Correlation (r)	True Odds Ratio			
	TCM at sampling point #11 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
TTHM at site 11	1.00	1.20	2.00	3.49	6.97
BDCM at site 11	0.81	1.16	1.75	2.76	4.83
THAA at site 11	0.26	1.05	1.20	1.39	1.66
DCAA at site 11	0.32	1.06	1.25	1.49	1.86
TCAA at site 11	0.09	1.02	1.06	1.12	1.19

	Correlation (r)	True Odds Ratio			
	TCM at sampling point #12 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
TTHM at site 12	1.00	1.20	2.00	3.50	6.99
BDCM at site 12	0.87	1.17	1.83	2.99	5.47
THAA at site 12	0.35	1.07	1.27	1.55	1.98
DCAA at site 12	0.37	1.07	1.29	1.59	2.05
TCAA at site 12	0.22	1.04	1.16	1.32	1.53

	Correlation (r)	True Odds Ratio			
	TCM at sampling point #13 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
TTHM at site 13	1.00	1.20	2.00	3.50	6.99
BDCM at site 13	0.87	1.17	1.83	2.99	5.48
THAA at site 13	0.18	1.03	1.13	1.25	1.42
DCAA at site 13	0.28	1.05	1.21	1.42	1.72
TCAA at site 13	-0.01	1.00	0.99	0.99	0.98

OR_T = the assumed OR for the true data

OR_X = the calculated OR for the measured data = OR_T^r

Sites 1-15 are in the distribution system

Table A6-9, continued. Correlations between TCM and other DBP species at sampling points #2 and #3 and other distribution system locations, and their effects on the odds ratio

City C Monthly

	Correlation (r)	True Odds Ratio			
	TCM at sampling point #14 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
TTHM at site 14	1.00	1.20	2.00	3.50	6.99
BDCM at site 14	0.89	1.18	1.85	3.03	5.61
THAA at site 14	0.15	1.03	1.11	1.21	1.34
DCAA at site 14	0.29	1.05	1.22	1.44	1.76
TCAA at site 14	-0.01	1.00	0.99	0.99	0.98

	Correlation (r)	True Odds Ratio			
	TCM at sampling point #15 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_X	OR_X	OR_X	OR_X
TTHM at site 15	1.00	1.20	2.00	3.50	6.99
BDCM at site 15	0.89	1.18	1.86	3.06	5.68
THAA at site 15	0.25	1.05	1.19	1.37	1.63
DCAA at site 15	0.21	1.04	1.16	1.30	1.50
TCAA at site 15	0.15	1.03	1.11	1.21	1.34

OR_T = the assumed OR for the true data

OR_X = the calculated OR for the measured data = OR_T^r

Sites 1-15 are in the distribution system

Table A6-10. Sensitivity and specificity between TCM and other DBP concentrations at Water treatment plant #1 and Water treatment plant #2 water treatment plants and distribution system locations, and their effects on the odds ratio

City A Daily

			True Odds Ratio			
	TCM at WTP1 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
DCAN at WTP1	0.86	0.95	1.15	1.75	2.78	4.84
DCAA at WTP1	0.75	0.85	1.10	1.46	2.04	3.07
TCAA at WTP1	0.86	0.95	1.10	1.75	2.78	4.84

			True Odds Ratio			
	TCM at WTP2 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
DCAN at WTP2	0.86	0.95	1.15	1.75	2.78	4.84
DCAA at WTP2	0.50	0.79	1.05	1.21	1.44	1.78
TCAA at WTP2	0.67	0.89	1.08	1.39	1.86	2.65

			True Odds Ratio			
	TCM at A01 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
DCAN at site A01	0.71	0.91	1.12	1.55	2.24	3.45
DCAA at site A01	0.57	0.86	1.08	1.35	1.77	2.42
TCAA at site A01	0.71	0.91	1.12	1.55	2.24	3.45

			True Odds Ratio			
	TCM at A02 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
DCAN at site A02	0.86	0.95	1.15	1.75	2.78	4.84
DCAA at site A02	0.71	0.91	1.12	1.55	2.24	3.45
TCAA at site A02	0.86	0.95	1.15	1.75	2.78	4.84

			True Odds Ratio			
	TCM at A03 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
DCAN at site A03	0.86	0.95	1.15	1.75	2.78	4.84
DCAA at site A03	0.57	0.86	1.08	1.35	1.77	2.42
TCAA at site A03	0.43	0.82	1.04	1.20	1.41	1.72

			True Odds Ratio			
	TCM at A04 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
DCAN at site A04	0.86	0.95	1.15	1.75	2.78	4.84
DCAA at site A04	0.57	0.86	1.08	1.35	1.77	2.42
TCAA at site A04	0.86	0.95	1.15	1.75	2.78	2.84

OR_T = the assumed OR for the true data

OR_X = the calculated OR for the measured data

WTP1=Water treatment plant #1 reservoir; WTP2=Water treatment plant #2 reservoir

Sites A01-A04 are home sampling locations

P_N = 0.2; P_D (OR_T=1.20)=0.23; P_D (OR_T=2.00)=0.33; P_D (OR_T=3.50)=0.0.47; P_D (OR_T=7.00)=0.64;

Table A6-11. Sensitivity and specificity between TCM and other DBP concentrations at Water treatment plant water treatment plant and distribution system locations, and their effects on the odds ratio

City B Daily

			True Odds Ratio			
	TCM at WTP (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
BDCM at WTP	0.57	0.86	1.08	1.35	1.77	2.42
DCAA at WTP	0.43	0.82	1.04	1.20	1.41	1.72
TCAA at WTP	0.29	0.77	1.01	1.04	1.09	1.15

			True Odds Ratio			
	TCM at B01 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
BDCM at site B01	0.57	0.86	1.08	1.35	1.77	2.42
DCAA at site B01	0.57	0.86	1.08	1.35	1.77	2.42
TCAA at site B01	0.57	0.86	1.08	1.35	1.77	2.42

			True Odds Ratio			
	TCM at B02 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
BDCM at site B02	0.33	0.78	1.02	1.08	1.17	1.28
DCAA at site B02	0.43	0.82	1.04	1.20	1.41	1.72
TCAA at site B02	0.29	0.77	1.01	1.04	1.09	1.15

			True Odds Ratio			
	TCM at B03 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
BDCM at site B03	0.00	0.00	1.20	2.00	3.50	7.00
DCAA at site B03	0.43	0.82	1.04	1.20	1.41	1.72
TCAA at site B03	0.71	0.91	1.12	1.55	2.24	3.45

			True Odds Ratio			
	TCM at B04 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
BDCM at site B04	0.43	0.81	1.04	1.18	1.39	1.67
DCAA at site B04	0.43	0.81	1.04	1.18	1.39	1.67
TCAA at site B04	0.29	0.76	1.01	1.04	1.07	1.12

OR_T = the assumed OR for the true data
OR_X = the calculated OR for the measured data
WTP=Water treatment plant reservoir
Sites B01-B04 are home sampling locations
P_N = 0.2; P_D (OR_T=1.20)=0.23; P_D (OR_T=2.00)=0.33; P_D (OR_T=3.50)=0.0.47; P_D (OR_T=7.00)=0.64
NC = Not calculable

Table A6-12. Sensitivity and specificity between TCM or TTHM and other DBP concentrations at Water treatment plant #1 and Water treatment plant #2 water treatment plants and distribution system locations, and their effects on the odds ratio

City A Monthly

			True Odds Ratio			
	TCM at WTP1 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
TTHM at WTP1	0.63	1.00	1.18	1.84	2.89	4.64
BDCM at WTP1	0.83	0.94	1.15	1.72	2.64	4.23
DCAA at WTP1	0.69	0.97	1.15	1.75	2.65	4.35
TCAA at WTP1	0.73	0.92	1.13	1.60	2.37	3.77
THAA at WTP1	0.75	0.96	1.11	1.60	2.11	3.09

			True Odds Ratio			
	TCM at WTP2 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
TTHM at WTP2	0.71	1.00	1.11	1.57	2.42	4.41
BDCM at WTP2	0.75	0.92	1.09	1.45	2.06	3.22
DCAA at WTP2	0.55	0.92	1.11	1.54	2.30	3.90
TCAA at WTP2	0.60	0.92	1.13	1.65	2.53	4.21
THAA at WTP2	0.50	0.93	0.89	0.72	0.64	0.60

			True Odds Ratio			
	TTHM at WTP1 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
TCM at WTP1	1.00	0.81	1.15	1.70	2.65	4.49
BDCM at WTP1	1.00	0.81	1.12	1.59	2.34	3.67
DCAA at WTP1	0.85	0.80	1.15	1.87	2.99	4.98
TCAA at WTP1	0.91	0.75	1.10	1.47	2.03	2.90
THAA at WTP1	0.92	0.83	1.10	1.47	2.01	2.83

			True Odds Ratio			
	TTHM at WTP2 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
TCM at WTP2	0.00	0.86	1.11	1.57	2.42	4.41
BDCM at WTP2	0.83	0.92	1.09	1.43	2.04	3.28
DCAA at WTP2	0.82	0.77	1.13	1.64	2.60	4.85
TCAA at WTP2	0.90	0.78	1.08	1.39	1.91	2.88
THAA at WTP2	0.83	0.80	1.09	1.44	2.02	3.11

OR_T = the assumed OR for the true data
OR_X = the calculated OR for the measured data
WTP1=Water treatment plant #1 reservoir; WTP2=Water treatment plant #2 reservoir
P_N = 0.2; P_D (OR_T=1.20)=0.23; P_D (OR_T=2.00)=0.33; P_D (OR_T=3.50)=0.0.47; P_D (OR_T=7.00)=0.64;
NC = Not calculable

Table A6-13. Sensitivity and specificity between monthly TCM and other DBP concentrations at water treatment plant and distribution system locations, and their effects on the odds ratio

City B Monthly

			True Odds Ratio			
	TCM at WTP (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
BDCM at WTP	0.78	0.93	1.13	1.64	2.49	4.06
DCAA at WTP	0.40	0.88	1.06	1.27	1.57	1.99
TCAA at WTP	0.40	0.88	1.06	1.27	1.57	1.99
THAA at WTP	0.40	0.88	1.06	1.27	1.57	1.99

			True Odds Ratio			
	TCM at B01 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
BDCM at site B01	0.50	0.85	1.06	1.29	1.62	2.11
DCAA at site B01	0.50	0.83	1.06	1.26	1.55	1.98
TCAA at site B01	0.75	0.92	1.13	1.60	2.37	3.77
THAA at site B01	0.75	0.92	1.13	1.60	2.37	3.77

			True Odds Ratio			
	TCM at B02 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
BDCM at site B02	0.75	0.93	1.13	1.63	2.44	3.90
DCAA at site B02	0.75	0.93	1.13	1.63	2.44	3.90
TCAA at site B02	0.50	0.83	1.06	1.26	1.55	1.98
THAA at site B02	0.50	0.83	1.06	1.26	1.55	1.98

			True Odds Ratio			
	TCM at B03 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
BDCM at site B03	0.75	0.92	1.13	1.60	2.37	3.77
DCAA at site B03	0.50	0.95	1.10	1.44	1.94	2.71
TCAA at site B03	0.60	0.91	1.10	1.48	2.05	2.98
THAA at site B03	0.50	0.83	1.06	1.26	1.55	1.98

			True Odds Ratio			
	TCM at B04 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
BDCM at site B04	0.75	0.85	1.06	1.29	1.62	2.11
DCAA at site B04	0.50	0.91	1.12	1.57	2.32	3.65
TCAA at site B04	0.60	0.90	1.10	1.45	2.00	2.88
THAA at site B04	0.50	0.90	1.10	1.45	2.00	2.88

OR_T = the assumed OR for the true data; OR_X = the calculated OR for the measured data

WTP=Water treatment plant reservoir; Sites B01-B04 are in the distribution system

P_N = 0.2; P_D (OR_T=1.20)=0.23; P_D (OR_T=2.00)=0.33; P_D (OR_T=3.50)=0.0.47; P_D (OR_T=7.00)=0.64

NC = Not calculable

TTHM sensitivity and specificity both equal 1 and result in no OR attenuation. Therefore these values are not included

Table A6-14. Sensitivity and specificity between TTHM and other DBP concentrations at Water treatment plant water treatment plant and distribution system locations, and their effects on the odds ratio

City B Monthly

			True Odds Ratio			
	TTHM at WTP (Measured)		$OR_T=1.20$	$OR_T=2.00$	$OR_T=3.50$	$OR_T=7.00$
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
BDCM at WTP	0.75	0.83	1.09	1.42	1.97	2.92
DCAA at WTP	0.80	0.81	1.09	1.43	1.99	3.02
TCAA at WTP	0.80	0.81	1.09	1.43	1.99	3.02
THAA at WTP	0.80	0.82	1.09	1.44	2.02	3.09

			True Odds Ratio			
	TTHM at B01 (Measured)		$OR_T=1.20$	$OR_T=2.00$	$OR_T=3.50$	$OR_T=7.00$
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
BDCM at site B01	0.75	0.85	1.10	1.46	2.04	3.07
DCAA at site B01	0.83	0.89	1.12	1.58	2.36	3.86
TCAA at site B01	0.83	0.89	1.12	1.58	2.36	3.86
THAA at site B01	0.83	0.90	1.12	1.60	2.41	3.97

			True Odds Ratio			
	TTHM at B02 (Measured)		$OR_T=1.20$	$OR_T=2.00$	$OR_T=3.50$	$OR_T=7.00$
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
BDCM at site B02	1.00	0.83	1.12	1.60	2.49	4.57
DCAA at site B02	0.67	0.89	1.10	1.48	2.07	3.06
TCAA at site B02	0.67	0.89	1.10	1.48	2.07	3.06
THAA at site B02	0.67	0.89	1.10	1.48	2.07	3.06

			True Odds Ratio			
	TTHM at B03 (Measured)		$OR_T=1.20$	$OR_T=2.00$	$OR_T=3.50$	$OR_T=7.00$
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
BDCM at site B03	0.67	0.69	1.05	1.22	1.48	1.89
DCAA at site B03	0.67	0.89	1.10	1.48	1.07	3.06
TCAA at site B03	0.80	0.89	1.12	1.56	2.31	3.69
THAA at site B03	0.67	0.89	1.10	1.48	2.07	3.06

			True Odds Ratio			
	TTHM at B04 (Measured)		$OR_T=1.20$	$OR_T=2.00$	$OR_T=3.50$	$OR_T=7.00$
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
BDCM at site B04	1.00	0.77	1.10	1.52	2.30	4.13
DCAA at site B04	0.80	0.89	1.12	1.56	2.31	3.69
TCAA at site B04	0.83	0.94	1.15	1.70	2.65	4.49
THAA at site B04	0.78	0.83	1.09	1.44	2.02	3.06

OR_T = the assumed OR for the true data; OR_X = the calculated OR for the measured data

WTP=Water treatment plant reservoir

Sites B01-B04 are in the distribution system

$P_N = 0.2$; P_D (OR_T=1.20)=0.23; P_D (OR_T=2.00)=0.33; P_D (OR_T=3.50)=0.0.47; P_D (OR_T=7.00)=0.64

NC = Not calculable

TCM sensitivity and specificity both equal 1 and result in no OR attenuation. Therefore these values are not included

Table A6-15. Sensitivity and specificity between monthly TCM and other DBP concentrations at distribution system locations, and their effects on the odds ratio

City C Monthly

	TCM at sampling point #2 (Measured)		True Odds Ratio			
			OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
BDCM2	0.80	0.94	1.14	1.69	2.60	4.31
THAA2	0.67	1.00	1.18	1.86	2.94	4.79
DCAA2	0.80	1.00	1.19	1.91	3.13	5.44
TCAA2	0.60	0.93	1.11	1.53	2.18	3.22

	TCM at sampling point #3 (Measured)		True Odds Ratio			
			OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
BDCM3	1.00	0.80	1.11	1.56	2.34	4.33
THAA3	0.00	0.60	0.94	0.77	0.58	0.36
DCAA3	0.00	0.60	0.94	0.77	0.58	0.36
TCAA3	0.00	0.60	0.94	0.77	0.58	0.36

	TCM at sampling point #4 (Measured)		True Odds Ratio			
			OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
BDCM4	1.00	1.00	1.20	2.00	3.50	7.00
THAA4	0.00	0.77	0.95	0.81	0.62	0.40
DCAA4	0.50	0.75	1.04	1.17	1.35	1.62
TCAA4	0.00	0.73	0.95	0.80	0.61	0.40

	TCM at sampling point #5 (Measured)		True Odds Ratio			
			OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
BDCM5	0.80	0.93	1.14	1.66	2.53	4.17
THAA5	0.00	0.69	0.95	0.79	0.60	0.39
DCAA5	0.25	0.69	0.99	0.96	0.93	0.88
TCAA5	0.00	0.69	0.95	0.79	0.60	0.39

OR_T = the assumed OR for the true data
OR_X = the calculated OR for the measured data
Sites 1-15 are in the distribution system
P_N = 0.2; P_D (OR_T=1.20)=0.23; P_D (OR_T=2.00)=0.33; P_D (OR_T=3.50)=0.0.47; P_D (OR_T=7.00)=0.64
NC = Not calculable
TTHM sensitivity and specificity both equal 1 and result in no OR attenuation. Therefore these values are not included

Table A6-15, continued. Sensitivity and specificity between monthly TCM and other DBP concentrations at distribution system locations, and their effects on the odds ratio

City C Monthly

			True Odds Ratio			
	TCM at sampling point #6 (Measured)		$OR_T=1.20$	$OR_T=2.00$	$OR_T=3.50$	$OR_T=7.00$
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
BDCM6	1.00	1.00	1.20	2.00	3.50	7.00
THAA6	0.00	0.69	0.95	0.79	0.60	0.39
DCAA6	0.00	0.69	0.95	0.79	0.60	0.39
TCAA6	0.00	0.67	0.95	0.79	0.60	0.38

			True Odds Ratio			
	TCM at sampling point #7 (Measured)		$OR_T=1.20$	$OR_T=2.00$	$OR_T=3.50$	$OR_T=7.00$
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
BDCM7	0.67	0.88	1.10	0.46	2.02	2.97
THAA7	0.25	0.71	0.99	0.97	0.95	0.92
DCAA7	0.40	0.75	1.02	1.10	1.21	1.36
TCAA7	0.25	0.69	0.99	0.96	0.93	0.88

			True Odds Ratio			
	TCM at sampling point #8 (Measured)		$OR_T=1.20$	$OR_T=2.00$	$OR_T=3.50$	$OR_T=7.00$
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
BDCM8	0.67	0.88	1.10	1.46	2.02	2.97
THAA8	0.67	0.82	1.08	1.36	1.79	2.51
DCAA8	0.60	0.87	1.08	1.39	1.86	2.60
TCAA8	0.40	0.80	1.03	1.15	1.31	1.54

			True Odds Ratio			
	TCM at sampling point #9 (Measured)		$OR_T=1.20$	$OR_T=2.00$	$OR_T=3.50$	$OR_T=7.00$
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
BDCM9	0.83	0.94	1.15	1.70	2.65	4.49
THAA9	0.67	0.79	1.07	1.20	1.71	2.34
DCAA9	0.60	0.82	1.07	1.31	1.68	2.26
TCAA9	0.40	0.76	1.03	1.11	1.23	1.39

OR_T = the assumed OR for the true data
OR_X = the calculated OR for the measured data
Sites 1-15 are in the distribution system
$P_N = 0.2$; P_D (OR_T=1.20)=0.23; P_D (OR_T=2.00)=0.33; P_D (OR_T=3.50)=0.0.47; P_D (OR_T=7.00)=0.64
NC = Not calculable
TTHM sensitivity and specificity both equal 1 and result in no OR attenuation. Therefore these values are not included

Table A6-15, continued. Sensitivity and specificity between monthly TCM and other DBP concentrations at distribution system locations, and their effects on the odds ratio

City C Monthly

			True Odds Ratio			
	TCM at sampling point #10 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
BDCM10	0.67	0.89	1.10	1.48	2.07	3.06
THAA10	0.67	0.79	1.08	1.32	1.71	2.34
DCAA10	0.60	0.82	1.07	1.31	1.68	2.26
TCAA10	0.40	0.76	1.03	1.11	1.23	1.39

			True Odds Ratio			
	TCM at sampling point #11 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
BDCM11	0.67	0.89	1.10	1.48	2.07	3.06
THAA11	0.40	0.76	1.03	1.11	1.23	1.39
DCAA11	0.50	0.81	1.05	1.23	1.49	1.88
TCAA11	0.60	0.82	1.07	1.31	1.68	2.26

			True Odds Ratio			
	TCM at sampling point #12 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
BDCM12	0.67	0.89	1.10	1.48	2.07	3.06
THAA12	0.33	0.73	1.01	1.04	1.08	1.13
DCAA12	0.40	0.75	1.02	1.10	1.21	1.36
TCAA12	0.20	0.69	0.98	0.93	0.86	0.78

			True Odds Ratio			
	TCM at sampling point #13 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
BDCM13	0.67	0.89	1.10	0.48	2.07	3.06
THAA13	0.50	0.81	1.15	1.23	1.49	1.88
DCAA13	0.40	0.76	1.03	1.11	1.23	1.39
TCAA13	0.20	0.71	0.99	0.94	0.88	0.81

OR_T = the assumed OR for the true data

OR_X = the calculated OR for the measured data

Sites 1-15 are in the distribution system

P_N = 0.2; P_D (OR_T=1.20)=0.23; P_D (OR_T=2.00)=0.33; P_D (OR_T=3.50)=0.0.47; P_D (OR_T=7.00)=0.64

NC = Not calculable

TTHM sensitivity and specificity both equal 1 and result in no OR attenuation. Therefore these values are not included

Table A6-15, continued. Sensitivity and specificity between monthly TCM and other DBP concentrations at distribution system locations, and their effects on the odds ratio

City C Monthly						
			True Odds Ratio			
	TCM at sampling point #14 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
BDCM14	0.67	0.89	1.10	1.48	2.07	3.06
THAA14	0.33	0.73	1.01	1.04	1.08	1.13
DCAA14	0.60	0.81	1.06	1.30	1.65	2.20
TCAA14	0.40	0.75	1.02	1.10	1.21	1.36
			True Odds Ratio			
	TCM at sampling point #15 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
BDCM15	0.67	0.88	1.10	1.46	2.02	2.97
THAA15	0.50	0.76	1.04	1.18	1.37	1.65
DCAA15	0.60	0.81	1.06	1.30	1.65	2.20
TCAA15	0.40	0.75	1.02	1.10	1.21	1.36

OR_T = the assumed OR for the true data
OR_X = the calculated OR for the measured data
Sites 1-15 are in the distribution system
P_N = 0.2; P_D (OR_T=1.20)=0.23; P_D (OR_T=2.00)=0.33; P_D (OR_T=3.50)=0.0.47; P_D (OR_T=7.00)=0.64
NC = Not calculable
TTHM sensitivity and specificity both equal 1 and result in no OR attenuation. Therefore these values are not included

Table A6-16. Sensitivity and specificity between monthly TTHM and other DBP concentrations at distribution system locations, and their effects on the odds ratio

City C Monthly

			True Odds Ratio			
	TTHM at sampling point #2 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
BDCM2	0.80	0.88	1.11	1.54	2.26	3.59
THAA2	0.96	0.89	1.13	1.67	2.64	4.79
DCAA2	1.00	1.00	1.20	2.00	3.50	7.00
TCAA2	0.80	0.93	1.12	1.66	2.53	4.17

			True Odds Ratio			
	TTHM at sampling point #3 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
BDCM3	1.00	0.60	1.08	1.38	1.96	3.31
THAA3	0.61	0.83	1.09	1.33	1.73	2.35
DCAA3	1.00	0.60	1.08	1.38	1.96	3.31
TCAA3	0.50	0.60	1.01	1.06	1.11	1.19

			True Odds Ratio			
	TTHM at sampling point #4 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
BDCM4	1.00	0.92	1.15	1.76	2.89	2.55
THAA4	0.76	0.71	1.06	1.29	1.67	2.30
DCAA4	0.50	0.63	1.02	1.07	1.15	1.26
TCAA4	0.00	0.64	0.95	0.78	0.59	0.37

			True Odds Ratio			
	TTHM at sampling point #5 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
BDCM5	1.00	0.80	1.11	1.56	2.39	4.33
THAA5	0.33	0.62	0.99	0.97	0.94	0.91
DCAA5	0.50	0.54	1.00	1.02	1.04	1.07
TCAA5	0.25	0.50	0.97	0.87	0.76	0.63

OR_T = the assumed OR for the true data

OR_X = the calculated OR for the measured data

Sites 1-15 are in the distribution system

P_N = 0.2; P_D (OR_T=1.20)=0.23; P_D (OR_T=2.00)=0.33; P_D (OR_T=3.50)=0.0.47; P_D (OR_T=7.00)=0.64

NC = Not calculable

TCM sensitivity and specificity both equal 1 and result in no OR attenuation. Therefore these values are not included

Table A6-16, continued. Sensitivity and specificity between monthly TTHM and other DBP concentrations at distribution system locations, and their effects on the odds ratio

City C Monthly

			True Odds Ratio			
	TTHM at sampling point #6 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
BDCM6	1.00	1.00	1.20	2.00	3.50	7.00
THAA6	0.11	0.68	0.97	0.87	0.74	0.59
DCAA6	1.00	0.69	0.19	1.45	2.12	3.68
TCAA6	0.00	0.75	0.95	0.80	0.62	0.40

			True Odds Ratio			
	TTHM at sampling point #7 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
BDCM7	1.00	0.76	1.10	1.51	2.28	4.06
THAA7	0.36	0.65	1.00	1.01	1.01	1.02
DCAA7	0.60	0.56	1.02	1.09	1.19	1.32
TCAA7	0.60	0.56	1.02	1.09	1.19	1.32

			True Odds Ratio			
	TTHM at sampling point #8 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
BDCM8	1.00	0.71	1.09	1.46	2.16	3.78
THAA8	0.36	0.67	1.00	1.02	0.34	1.06
DCAA8	0.60	0.60	1.02	1.11	1.23	1.39
TCAA8	0.40	0.53	0.98	0.96	0.93	0.88

			True Odds Ratio			
	TTHM at sampling point #9 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
BDCM9	0.00	0.72	1.09	1.47	2.18	3.83
THAA9	0.36	0.65	1.00	1.01	1.02	1.04
DCAA9	0.60	0.59	1.02	1.11	1.23	1.39
TCAA9	0.60	0.59	1.02	1.11	1.23	1.39

OR_T = the assumed OR for the true data
OR_X = the calculated OR for the measured data
Sites 1-15 are in the distribution system
P_N = 0.2; P_D (OR_T=1.20)=0.23; P_D (OR_T=2.00)=0.33; P_D (OR_T=3.50)=0.0.47; P_D (OR_T=7.00)=0.64
NC = Not calculable
TCM sensitivity and specificity both equal 1 and result in no OR attenuation. Therefore these values are not included

Table A6-16, continued. Sensitivity and specificity between monthly TTHM and other DBP concentrations at distribution system locations, and their effects on the odds ratio

City C Monthly

			True Odds Ratio			
	TTHM at sampling point #10 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
BDCM10	1.00	0.72	1.09	1.47	2.18	3.83
THAA10	0.27	0.61	0.98	0.93	0.87	0.79
DCAA10	0.60	0.59	1.05	1.23	1.49	1.88
TCAA10	0.40	0.53	0.99	0.96	0.93	0.88

			True Odds Ratio			
	TTHM at sampling point #11 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
BDCM11	0.67	0.89	1.10	1.48	2.07	3.06
THAA11	0.33	0.78	1.02	1.08	1.17	1.28
DCAA11	0.50	0.81	1.07	1.31	1.68	2.26
TCAA11	0.60	0.82	1.07	1.31	1.68	2.26

			True Odds Ratio			
	TTHM at sampling point #12 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
BDCM12	0.67	0.89	1.10	1.48	2.07	3.06
THAA12	0.33	0.76	1.01	1.06	1.13	1.22
DCAA12	0.40	0.75	1.02	1.10	1.21	1.36
TCAA12	0.20	0.69	0.98	0.93	0.86	0.78

			True Odds Ratio			
	TTHM at sampling point #13 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
BDCM13	0.67	0.89	1.10	1.48	2.07	3.06
THAA13	0.50	0.82	1.05	1.25	1.52	1.93
DCAA13	0.40	0.76	1.03	1.11	1.23	1.39
TCAA13	0.20	0.71	0.99	0.94	0.88	0.81

OR_T = the assumed OR for the true data

OR_X = the calculated OR for the measured data

Sites 1-15 are in the distribution system

P_N = 0.2; P_D (OR_T=1.20)=0.23; P_D (OR_T=2.00)=0.33; P_D (OR_T=3.50)=0.0.47; P_D (OR_T=7.00)=0.64

NC = Not calculable

TCM sensitivity and specificity both equal 1 and result in no OR attenuation. Therefore these values are not included

Table A6-16, continued. Sensitivity and specificity between monthly TTHM and other DBP concentrations at distribution system locations, and their effects on the odds ratio

City C Monthly

			True Odds Ratio			
	TTHM at sampling point #14 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
BDCM14	.67	0.89	1.10	1.48	2.07	3.06
THAA14	.33	0.76	1.01	1.06	1.13	1.22
DCAA14	.60	0.81	1.06	1.30	1.65	2.20
TCAA14	.40	0.75	1.02	1.10	1.21	1.36

			True Odds Ratio			
	TTHM at sampling point #15(Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True"	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
BDCM15	1.00	0.76	1.10	1.51	2.28	4.06
THAA15	0.27	0.70	1.00	0.98	0.96	0.94
DCAA15	0.60	0.63	1.03	1.30	1.65	2.20
TCAA15	0.40	0.56	0.99	0.98	0.96	0.93

OR_T = the assumed OR for the true data
OR_X = the calculated OR for the measured data
Sites 1-15 are in the distribution system
P_N = 0.2; P_D (OR_T=1.20)=0.23; P_D (OR_T=2.00)=0.33; P_D (OR_T=3.50)=0.0.47; P_D (OR_T=7.00)=0.64
NC = Not calculable
TCM sensitivity and specificity both equal 1 and result in no OR attenuation. Therefore these values are not included

Chapter 7

Table A7-1. Water treatment plant #1 TCM data; correlations between the "measured" convenience sampling data and the "true" data with the resulting OR attenuation due to exposure misclassification

Water treatment plant #1 TCM data

	Correlation (r)	True Odds Ratio			
"True" data	Convenience sampling #1 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_{Xa}	OR_{Xa}	OR_{Xa}	OR_{Xa}
Maximum TCM	0.89	1.18	1.86	3.06	5.69
Minimum TCM	0.89	1.18	1.85	3.05	5.66
Mean TCM	0.93	1.19	1.91	3.23	6.17
	Correlation (r)	True Odds Ratio			
"True" data	Convenience sampling #2 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_{Xa}	OR_{Xa}	OR_{Xa}	OR_{Xa}
Maximum TCM	0.90	1.18	1.86	3.08	5.74
Minimum TCM	0.92	1.18	1.89	3.17	6.01
Mean TCM	0.95	1.19	1.93	3.27	6.30
	Correlation (r)	True Odds Ratio			
"True" data	Convenience sampling #3 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_{Xa}	OR_{Xa}	OR_{Xa}	OR_{Xa}
Maximum TCM	0.87	1.17	1.83	2.97	5.43
Minimum TCM	0.82	1.16	1.77	2.80	4.95
Mean TCM	0.87	1.17	1.83	2.98	5.46
	Correlation (r)	True Odds Ratio			
"True" data	Convenience sampling #4 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_{Xa}	OR_{Xa}	OR_{Xa}	OR_{Xa}
Maximum TCM	0.88	1.17	1.84	3.00	5.51
Minimum TCM	0.84	1.16	1.78	2.85	5.08
Mean TCM	0.88	1.17	1.84	3.00	5.52

OR_T = the assumed OR for the true data

OR_X = the calculated OR for the measured data = $OR_T{}^r$

WTP1 = Water treatment plant #1; WTP2 = Water treatment plant #2

Table A7-2. Water treatment plant #2 TCM data; correlations between the "measured" convenience sampling data and the "true" data with the resulting OR attenuation due to exposure misclassification

Water treatment plant #2 TCM data

	Correlation (r)	True Odds Ratio			
"True" data	Convenience sampling #1 (Measured)	OR_T=1.20 OR_{Xa}	OR_T=2.00 OR_{Xa}	OR_T=3.50 OR_{Xa}	OR_T=7.00 OR_{Xa}
Maximum TCM	0.90	1.18	1.86	3.07	5.72
Minimum TCM	0.89	1.18	1.85	3.03	5.60
Mean TCM	0.94	1.19	1.92	3.25	6.23
	Correlation (r)	True Odds Ratio			
"True" data	Convenience sampling #2 (Measured)	OR_T=1.20 OR_{Xa}	OR_T=2.00 OR_{Xa}	OR_T=3.50 OR_{Xa}	OR_T=7.00 OR_{Xa}
Maximum TCM	0.94	1.19	1.92	3.26	6.26
Minimum TCM	0.88	1.17	1.84	3.00	5.51
Mean TCM	0.96	1.19	1.95	3.34	6.50
	Correlation (r)	True Odds Ratio			
"True" data	Convenience sampling #3 (Measured)	OR_T=1.20 OR_{Xa}	OR_T=2.00 OR_{Xa}	OR_T=3.50 OR_{Xa}	OR_T=7.00 OR_{Xa}
Maximum TCM	0.90	1.18	1.87	3.10	5.79
Minimum TCM	0.90	1.18	1.86	3.08	5.75
Mean TCM	0.95	1.19	1.94	3.30	6.39
	Correlation (r)	True Odds Ratio			
"True" data	Convenience sampling #4 (Measured)	OR_T=1.20 OR_{Xa}	OR_T=2.00 OR_{Xa}	OR_T=3.50 OR_{Xa}	OR_T=7.00 OR_{Xa}
Maximum TCM	0.90	1.18	1.87	3.10	5.80
Minimum TCM	0.90	1.18	1.86	3.08	5.73
Mean TCM	0.95	1.19	1.93	3.29	6.37

OR_T = the assumed OR for the true data

OR_X = the calculated OR for the measured data = OR_T^r

WTP1 = Water treatment plant #1; WTP2 = Water treatment plant #2

Table A7-3. Water treatment plant #1 TTHM data; correlations between the "measured" convenience sampling data and the "true" data with the resulting OR attenuation due to exposure misclassification

Water treatment plant #1 TTHM data					
	Correlation (r)	True Odds Ratio			
"True" data	Convenience sampling #1 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_{Xa}	OR_{Xa}	OR_{Xa}	OR_{Xa}
Maximum TTHM	0.89	1.18	1.86	3.07	5.70
Minimum TTHM	0.89	1.18	1.85	3.04	5.61
Mean TTHM	0.93	1.19	1.91	3.22	6.14
	Correlation (r)	True Odds Ratio			
"True" data	Convenience sampling #2 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_{Xa}	OR_{Xa}	OR_{Xa}	OR_{Xa}
Maximum TTHM	0.90	1.18	1.86	3.08	5.74
Minimum TTHM	0.92	1.18	1.89	3.16	5.98
Mean TTHM	0.95	1.19	1.93	3.29	6.35
	Correlation (r)	True Odds Ratio			
"True" data	Convenience sampling #3 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_{Xa}	OR_{Xa}	OR_{Xa}	OR_{Xa}
Maximum TTHM	0.87	1.17	1.83	2.97	5.43
Minimum TTHM	0.81	1.16	1.75	2.76	4.85
Mean TTHM	0.87	1.17	1.83	2.97	5.43
	Correlation (r)	True Odds Ratio			
"True" data	Convenience sampling #4 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_{Xa}	OR_{Xa}	OR_{Xa}	OR_{Xa}
Maximum TTHM	0.88	1.17	1.84	3.00	5.51
Minimum TTHM	0.83	1.16	1.78	2.83	5.03
Mean TTHM	0.88	1.17	1.84	3.01	5.53

OR_T = the assumed OR for the true data

OR_X = the calculated OR for the measured data = $OR_T^{\,r}$

WTP1 = Water treatment plant #1; WTP2 = Water treatment plant #2

Table A7-4. Water treatment plant #2 TTHM data; correlations between the "measured" convenience sampling data and the "true" data with the resulting OR attenuation due to exposure misclassification

Water treatment plant #2 TTHM data

	Correlation (r)	True Odds Ratio			
"True" data	Convenience sampling #1 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_{Xa}	OR_{Xa}	OR_{Xa}	OR_{Xa}
Maximum TTHM	0.89	1.18	1.86	3.06	5.68
Minimum TTHM	0.91	1.18	1.88	3.13	5.87
Mean TTHM	0.94	1.19	1.92	3.24	6.20
	Correlation (r)	True Odds Ratio			
"True" data	Convenience sampling #2 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_{Xa}	OR_{Xa}	OR_{Xa}	OR_{Xa}
Maximum TTHM	0.94	1.19	1.92	3.26	6.26
Minimum TTHM	0.91	1.18	1.89	3.15	5.93
Mean TTHM	0.96	1.19	1.95	3.34	6.51
	Correlation (r)	True Odds Ratio			
"True" data	Convenience sampling #3 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_{Xa}	OR_{Xa}	OR_{Xa}	OR_{Xa}
Maximum TTHM	0.90	1.18	1.87	3.09	5.76
Minimum TTHM	0.93	1.18	1.90	3.19	6.06
Mean TTHM	0.95	1.19	1.93	3.29	6.37
	Correlation (r)	True Odds Ratio			
"True" data	Convenience sampling #4 (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
		OR_{Xa}	OR_{Xa}	OR_{Xa}	OR_{Xa}
Maximum TTHM	0.90	1.18	1.87	3.09	5.76
Minimum TTHM	0.93	1.18	1.90	3.20	6.08
Mean TTHM	0.95	1.19	1.93	3.29	6.35

OR_T = the assumed OR for the true data

OR_X = the calculated OR for the measured data = OR_T^r

WTP1 = Water treatment plant #1; WTP2 = Water treatment plant #2

Table A7-5. Water treatment plant #1 data; correlations between the "measured" mean DBP concentration data and the "true" data with the resulting OR attenuation due to exposure misclassification

Water treatment plant #1 TCM data

"True" data	Correlation (r) Mean TCM (Measured)	True Odds Ratio OR_T=1.20 OR_{Xa}	 OR_T=2.00 OR_{Xa}	 OR_T=3.50 OR_{Xa}	 OR_T=7.00 OR_{Xa}
Maximum TCM	0.96	1.19	1.94	3.32	6.45
Minimum TCM	0.96	1.19	1.94	3.31	6.42
Convenience #1	0.93	1.19	1.91	3.23	6.17
Convenience #2	0.95	1.19	1.93	3.27	6.30
Convenience #3	0.87	1.17	1.83	2.98	5.46
Convenience #4	0.88	1.17	1.84	3.00	5.52

Water treatment plant #1 TTHM data

"True" data	Correlation (r) Mean TTHM (Measured)	True Odds Ratio OR_T=1.20 OR_{Xa}	 OR_T=2.00 OR_{Xa}	 OR_T=3.50 OR_{Xa}	 OR_T=7.00 OR_{Xa}
Maximum TTHM	0.96	1.19	1.94	3.32	6.45
Minimum TTHM	0.96	1.19	1.95	3.34	6.51
Convenience #1	0.93	1.19	1.91	3.22	6.14
Convenience #2	0.95	1.19	1.93	3.29	6.35
Convenience #3	0.87	1.17	1.83	2.97	5.43
Convenience #4	0.88	1.17	1.84	3.01	5.53

OR_T = the assumed OR for the true data

OR_X = the calculated OR for the measured data = $OR_T^{\,r}$

WTP1 = Water treatment plant #1; WTP2 = Water treatment plant #2

Table A7-6. Water treatment plant #2 data; correlations between the "measured" mean DBP concentration data and the "true" data with the resulting OR attenuation due to exposure misclassification

Water treatment plant #2 TCM data

	Correlation (r)	True Odds Ratio			
	Mean TCM (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True" data		OR_{Xa}	OR_{Xa}	OR_{Xa}	OR_{Xa}
Maximum TCM	0.97	1.19	1.96	3.38	6.64
Minimum TCM	0.93	1.19	1.91	3.21	6.13
Convenience #1	0.94	1.19	1.92	3.25	6.23
Convenience #2	0.96	1.19	1.95	3.34	6.50
Convenience #3	0.95	1.19	1.94	3.30	6.39
Convenience #4	0.95	1.19	1.93	3.29	6.37

Water treatment plant #2 TTHM data

	Correlation (r)	True Odds Ratio			
	Mean TTHM (Measured)	OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True" data		OR_{Xa}	OR_{Xa}	OR_{Xa}	OR_{Xa}
Maximum TTHM	0.97	1.19	1.96	3.38	6.63
Minimum TTHM	0.96	1.19	1.94	3.32	6.46
Convenience #1	0.94	1.19	1.92	3.24	6.20
Convenience #2	0.96	1.19	1.95	3.34	6.51
Convenience #3	0.95	1.19	1.93	3.29	6.37
Convenience #4	0.95	1.19	1.93	3.29	6.35

OR_T = the assumed OR for the true data
OR_X = the calculated OR for the measured data = OR_T'
WTP1 = Water treatment plant #1; WTP2 = Water treatment plant #2

Table A7-7. Categorical treatment of City A data comparing the "measured" convenience sampling data with the "true" values and the resulting effect on the OR

Water treatment plant #1 TCM						
			True Odds Ratio			
	Convenience #1 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True" data	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
Maximum TCM	0.75	0.92	1.13	1.60	2.37	3.77
Minimum TCM	0.92	0.97	1.17	1.85	3.06	5.63
Mean TCM	0.83	0.94	1.15	1.71	2.65	4.49
	Convenience #2 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True" data	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
Maximum TCM	0.83	0.94	1.15	1.70	2.65	4.49
Minimum TCM	0.83	0.94	1.15	1.70	2.65	4.49
Mean TCM	0.83	0.94	1.15	1.70	2.65	4.49
	Convenience #3 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True" data	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
Maximum TCM	0.83	0.94	1.15	1.70	2.65	4.49
Minimum TCM	0.92	0.97	1.17	1.85	3.06	5.63
Mean TCM	0.92	0.97	1.17	1.85	2.16	5.63
	Convenience #4 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True" data	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
Maximum TCM	0.83	0.94	1.15	1.70	2.65	4.49
Minimum TCM	0.92	0.97	1.17	1.85	3.06	5.63
Mean TCM	0.92	0.97	1.17	1.85	2.16	5.63

OR_T = the assumed OR for the true data

OR_X = the calculated OR for the measured data

WTP1=Water treatment plant #1 reservoir; WTP2=Water treatment plant #2 reservoir

P_N = 0.2; P_D (OR_T=1.20)=0.23; P_D (OR_T=2.00)=0.33; P_D (OR_T=3.50)=0.0.47; P_D (OR_T=7.00)=0.64;

Table A7-8. Categorical treatment of City A data comparing the "measured" convenience sampling data with the "true" values and the resulting effect on the OR

Water treatment plant #2 TCM						
			True Odds Ratio			
	Convenience #1 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True" data	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
Maximum TCM	0.83	0.94	1.15	1.70	2.65	4.49
Minimum TCM	0.83	0.94	1.15	1.70	2.65	4.49
Mean TCM	0.83	0.94	1.15	1.70	2.65	4.49
	Convenience #2 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True" data	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
Maximum TCM	0.92	0.97	1.17	1.85	3.06	5.63
Minimum TCM	0.67	0.89	1.10	1.48	2.07	3.06
Mean TCM	0.75	0.92	1.13	1.60	2.37	3.47
	Convenience #3 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True" data	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
Maximum TCM	0.75	0.92	1.13	1.60	2.37	3.77
Minimum TCM	0.92	0.97	1.17	1.85	3.06	5.63
Mean TCM	0.83	0.94	1.15	1.70	2.65	4.49
	Convenience #4 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True" data	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
Maximum TCM	0.83	0.94	1.15	1.70	2.66	4.50
Minimum TCM	0.75	0.94	1.15	1.70	2.66	4.51
Mean TCM	0.92	0.97	1.17	1.85	3.06	5.63

OR_T = the assumed OR for the true data

OR_X = the calculated OR for the measured data

WTP1=Water treatment plant #1 reservoir; WTP2=Water treatment plant #2 reservoir

P_N = 0.2; P_D (OR_T=1.20)=0.23; P_D (OR_T=2.00)=0.33; P_D (OR_T=3.50)=0.0.47; P_D (OR_T=7.00)=0.64;

Table A7-9. Categorical treatment of City A data comparing the "measured" convenience sampling data with the "true" values and the resulting effect on the OR

Water treatment plant #1 TTHM						
			True Odds Ratio			
	Convenience #1 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True" data	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
Maximum TTHM	0.75	0.92	1.13	1.60	2.37	3.77
Minimum TTHM	0.92	0.97	1.17	1.85	3.06	5.63
Mean TTHM	0.83	0.94	1.15	1.71	2.65	4.49
	Convenience #2 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True" data	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
Maximum TTHM	0.83	0.94	1.15	1.70	2.65	4.49
Minimum TTHM	0.83	0.94	1.15	1.70	2.65	4.49
Mean TTHM	0.83	0.94	1.15	1.70	2.65	4.49
	Convenience #3 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True" data	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
Maximum TTHM	0.83	0.94	1.19	1.92	3.18	5.62
Minimum TTHM	0.92	0.97	1.17	1.85	3.06	5.63
Mean TTHM	0.92	0.97	1.17	1.55	3.06	5.63
	Convenience #4 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True" data	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
Maximum TTHM	0.83	0.94	1.19	1.92	3.18	5.62
Minimum TTHM	0.92	0.97	1.17	1.85	3.06	5.63
Mean TTHM	0.92	0.97	1.17	1.55	3.06	5.63

OR_T = the assumed OR for the true data

OR_X = the calculated OR for the measured data

WTP1=Water treatment plant #1 reservoir; WTP2=Water treatment plant #2 reservoir

P_N = 0.2; P_D (OR_T=1.20)=0.23; P_D (OR_T=2.00)=0.33; P_D (OR_T=3.50)=0.0.47; P_D (OR_T=7.00)=0.64;

Table A7-10. Categorical treatment of City A data comparing the "measured" convenience sampling data with the "true" values and the resulting effect on the OR

Water treatment plant #2 TTHM

			True Odds ratio			
	Convenience #1 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True" data	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
Maximum TTHM	0.75	0.92	1.13	1.60	2.37	3.77
Minimum TTHM	0.75	0.92	1.13	1.60	2.37	3.77
Mean TTHM	0.67	0.89	1.10	1.48	2.07	3.06
	Convenience #2 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True" data	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
Maximum TTHM	0.83	0.94	1.15	1.70	2.65	4.49
Minimum TTHM	0.75	0.92	1.13	1.60	2.37	3.77
Mean TTHM	0.75	0.92	1.13	1.60	2.37	3.77
	Convenience #3 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True" data	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
Maximum TTHM	0.75	0.92	1.13	1.60	2.37	3.77
Minimum TTHM	0.83	0.94	1.15	1.70	22.65	4.49
Mean TTHM	0.92	0.97	1.17	1.85	3.06	5.63
	Convenience #4 (Measured)		OR_T=1.20	OR_T=2.00	OR_T=3.50	OR_T=7.00
"True" data	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
Maximum TTHM	0.83	0.94	1.15	1.70	2.65	4.49
Minimum TTHM	0.75	0.92	1.13	1.60	3.37	3.77
Mean TTHM	0.92	0.97	1.17	1.85	3.06	5.63

OR_T = the assumed OR for the true data

OR_X = the calculated OR for the measured data

WTP1=Water treatment plant #1 reservoir; WTP2=Water treatment plant #2 reservoir

P_N = 0.2; P_D (OR_T=1.20)=0.23; P_D (OR_T=2.00)=0.33; P_D (OR_T=3.50)=0.0.47; P_D (OR_T=7.00)=0.64;

Table A7-11. Categorical treatment of City A data comparing the "measured" mean DBP concentration data with the "true" values and the resulting effect on the OR

Water treatment plant #1 TCM

	Mean TCM (Measured)		$OR_T=1.20$	$OR_T=2.00$	$OR_T=3.50$	$OR_T=7.00$
"True" data	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
Maximum TCM	0.92	0.97	1.17	1.85	3.06	5.63
Minimum TCM	0.83	0.94	1.15	1.70	2.65	4.49

Water treatment plant #2 TCM

	Mean TCM (Measured)		$OR_T=1.20$	$OR_T=2.00$	$OR_T=3.50$	$OR_T=7.00$
"True" data	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
Maximum TCM	0.83	0.94	1.15	1.70	2.65	4.49
Minimum TCM	0.92	0.97	1.17	1.85	3.06	5.63

Water treatment plant #1 TTHM

	Mean TTHM (Measured)		$OR_T=1.20$	$OR_T=2.00$	$OR_T=3.50$	$OR_T=7.00$
"True" data	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
Maximum TTHM	0.92	0.97	1.17	1.85	3.06	5.63
Minimum TTHM	0.83	0.94	1.15	1.70	2.65	4.49

Water treatment plant #2 TTHM

	Mean TTHM (Measured)		$OR_T=1.20$	$OR_T=2.00$	$OR_T=3.50$	$OR_T=7.00$
"True" data	Sensitivity	Specificity	OR_X	OR_X	OR_X	OR_X
Maximum TTHM	0.83	0.94	1.15	1.70	2.65	4.49
Minimum TTHM	0.83	0.94	1.15	1.70	2.65	4.49

OR_T = the assumed OR for the true data

OR_X = the calculated OR for the measured data

WTP1=Water treatment plant #1 reservoir; WTP2=Water treatment plant #2 reservoir

$P_N = 0.2$; P_D $(OR_T=1.20)=0.23$; P_D $(OR_T=2.00)=0.33$; P_D $(OR_T=3.50)=0.0.47$; P_D $(OR_T=7.00)=0.64$;